PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/12, C07K 14/47, A61K 38/17 — A1

(11) International Publication Number: **WO 98/49302**

(43) International Publication Date: 5 November 1998 (05.11.98)

(21) International Application Number: PCT/US98/08336

(22) International Filing Date: 24 April 1998 (24.04.98)

(30) Priority Data:

08/845,296	25 April 1997 (25.04.97)	US
09/065,125	23 April 1998 (23.04.98)	US

(71) Applicant: GENETICS INSTITUTE, INC. [US/US]; 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(72) Inventors: JACOBS, Kenneth; 151 Beaumont Avenue, Newton, MA 02160 (US). MCCOY, John, M.; 56 Howard Street, Reading, MA 01867 (US). LAVALLIE, Edward, R.; 113 Ann Lee Road, Harvard, MA 01451 (US). RACIE, Lisa, A.; 124 School Street, Acton, MA 01720 (US). MERBERG, David; 2 Orchard Drive, Acton, MA 01720 (US). TREACY, Maurice; 93 Walcott Road, Chestnut Hill, MA 02167 (US). SPAULDING, Vikki; 11 Meadowbank Road, Billerica, MA 01821 (US). AGOSTINO, Michael, J.; 26 Wolcott Avenue, Andover, MA 01810 (US).

(74) Agent: SPRUNGER, Suzanne, A.; Genetics Institute, Inc., 87 CambridgePark Drive, Cambridge, MA 02140 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, GW, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(57) Abstract

Polynucleotides and the proteins encoded thereby are disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

This application is a continuation-in-part of Ser. No. 60/XXX,XXX (converted to a provisional application from non-provisional application Ser. No. 08/845,296), filed April 25, 1997, which is incorporated by reference herein.

FIELD OF THE INVENTION

The present invention provides novel polynucleotides and proteins encoded by such polynucleotides, along with therapeutic, diagnostic and research utilities for these polynucleotides and proteins.

BACKGROUND OF THE INVENTION

Technology aimed at the discovery of protein factors (including e.g., cytokines, such as lymphokines, interferons, CSFs and interleukins) has matured rapidly over the past decade. The now routine hybridization cloning and expression cloning techniques clone novel polynucleotides "directly" in the sense that they rely on information directly related to the discovered protein (i.e., partial DNA/amino acid sequence of the protein in the case of hybridization cloning; activity of the protein in the case of expression cloning). More recent "indirect" cloning techniques such as signal sequence cloning, which isolates DNA sequences based on the presence of a now well-recognized secretory leader sequence motif, as well as various PCR-based or low stringency hybridization cloning techniques, have advanced the state of the art by making available large numbers of DNA/amino acid sequences for proteins that are known to have biological activity by virtue of their secreted nature in the case of leader sequence cloning, or by virtue of the cell or tissue source in the case of PCR-based techniques. It is to these proteins and the polynucleotides encoding them that the present invention is directed.

SUMMARY OF THE INVENTION

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 99 to nucleotide 902;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 162 to nucleotide 902;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 87 to nucleotide 219;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ci25_4 deposited under accession number ATCC 98415;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ci25_4 deposited under accession number ATCC 98415;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ci25_4 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ci25_4 deposited under accession number ATCC 98415;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 129 to amino acid 138 of SEQ ID NO:2;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:1 from nucleotide 99 to nucleotide 902; the nucleotide sequence of SEQ ID NO:1 from nucleotide 162 to nucleotide 902; the nucleotide sequence of SEQ ID NO:1 from nucleotide 87 to nucleotide 219; the nucleotide sequence of the full-length protein coding sequence of clone ci25_4 deposited under accession number ATCC 98415; or the nucleotide sequence of a mature protein coding sequence of clone ci25_4 deposited under accession number ATCC 98415. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone ci25_4 deposited under accession number ATCC 98415.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:1.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 129 to amino acid 138 of SEQ ID NO:2; and

(c) the amino acid sequence encoded by the cDNA insert of clone ci25_4 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:2.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 283 to nucleotide 1158;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 789;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone da228_6 deposited under accession number ATCC 98415;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone da228_6 deposited under accession number ATCC 98415;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone da228_6 deposited under accession number ATCC 98415;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone da228_6 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 141 to amino acid 150 of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:3 from nucleotide 283 to nucleotide 1158; the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 789; the nucleotide sequence of the full-length protein coding sequence of clone da228_6 deposited under accession number ATCC 98415; or the nucleotide sequence of a mature protein coding sequence of clone da228_6 deposited under accession number ATCC 98415. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone da228_6 deposited under accession number ATCC 98415. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 169.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:3.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 169;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 141 to amino acid 150 of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone da228_6 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:4 or the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 169.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 152 to nucleotide 2182;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 2 to nucleotide 931;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone du410_5 deposited under accession number ATCC 98415;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone du410_5 deposited under accession number ATCC 98415;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone du410_5 deposited under accession number ATCC 98415;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone du410_5 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment

comprising the amino acid sequence from amino acid 333 to amino acid 342 of SEQ ID NO:6;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:5 from nucleotide 152 to nucleotide 2182; the nucleotide sequence of SEQ ID NO:5 from nucleotide 2 to nucleotide 931; the nucleotide sequence of the full-length protein coding sequence of clone du410_5 deposited under accession number ATCC 98415; or the nucleotide sequence of a mature protein coding sequence of clone du410_5 deposited under accession number ATCC 98415. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone du410_5 deposited under accession number ATCC 98415. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 260.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:5.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 260;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 333 to amino acid 342 of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone du410_5 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:6 or the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 260.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 51 to nucleotide 611;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 1 to nucleotide 525;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone eh80_1 deposited under accession number ATCC 98415;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone eh80_1 deposited under accession number ATCC 98415;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone eh80_1 deposited under accession number ATCC 98415;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone eh80_1 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 88 to amino acid 97 of SEQ ID NO:8;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:7 from nucleotide 51 to nucleotide 611; the nucleotide sequence of SEQ ID NO:7 from nucleotide 1 to nucleotide 525; the nucleotide sequence of the full-length protein coding sequence of clone eh80_1 deposited under accession number ATCC 98415; or the

nucleotide sequence of a mature protein coding sequence of clone eh80_1 deposited under accession number ATCC 98415. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone eh80_1 deposited under accession number ATCC 98415. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 158.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:7.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 158;

(c) fragments of the amino acid sequence of SEQ ID NO:8 comprising the amino acid sequence from amino acid 88 to amino acid 97 of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone eh80_1 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:8 or the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 158.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 431 to nucleotide 559;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 518 to nucleotide 559;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 190 to nucleotide 547;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone er369_1 deposited under accession number ATCC 98415;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone er369_1 deposited under accession number ATCC 98415;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone er369_1 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone er369_1 deposited under accession number ATCC 98415;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 16 to amino acid 25 of SEQ ID NO:10;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:9 from nucleotide 431 to nucleotide 559; the nucleotide sequence of SEQ ID NO:9 from nucleotide 518 to nucleotide 559; the nucleotide sequence of SEQ ID NO:9 from nucleotide 190 to nucleotide 547; the nucleotide sequence of the full-length protein coding sequence of clone er369_1 deposited under accession number ATCC 98415; or the nucleotide sequence of a mature protein coding sequence of clone er369_1 deposited under accession number ATCC 98415. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone er369_1 deposited under accession number ATCC 98415. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10 from amino acid 1 to amino acid 39.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:9.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:10 from amino acid 1 to amino acid 39;

(c) fragments of the amino acid sequence of SEQ ID NO:10 comprising the amino acid sequence from amino acid 16 to amino acid 25 of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone er369_1 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:10 or the amino acid sequence of SEQ ID NO:10 from amino acid 1 to amino acid 39.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 91 to nucleotide 2838;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 2209 to nucleotide 2838;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 839 to nucleotide 1197;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fh123_5 deposited under accession number ATCC 98415;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fh123_5 deposited under accession number ATCC 98415;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fh123_5 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fh123_5 deposited under accession number ATCC 98415;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 453 to amino acid 462 of SEQ ID NO:12;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:11 from nucleotide 91 to nucleotide 2838; the nucleotide sequence of SEQ ID NO:11 from nucleotide 2209 to nucleotide 2838; the nucleotide sequence of SEQ ID NO:11 from nucleotide 839 to nucleotide 1197; the nucleotide sequence of the full-length protein coding sequence of clone fh123_5 deposited under accession number ATCC 98415; or the nucleotide sequence of a mature protein coding sequence of clone fh123_5 deposited under accession number ATCC 98415. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fh123_5 deposited under accession number ATCC 98415. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12 from amino acid 251 to amino acid 369.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:11.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 251 to amino acid 369;

(c) fragments of the amino acid sequence of SEQ ID NO:12 comprising the amino acid sequence from amino acid 453 to amino acid 462 of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone fh123_5 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:12 or the amino acid sequence of SEQ ID NO:12 from amino acid 251 to amino acid 369.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 568 to nucleotide 978;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1084 to nucleotide 1854;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fm60_1 deposited under accession number ATCC 98415;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fm60_1 deposited under accession number ATCC 98415;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fm60_1 deposited under accession number ATCC 98415;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fm60_1 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising the amino acid sequence from amino acid 63 to amino acid 72 of SEQ ID NO:14;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:13 from nucleotide 568 to nucleotide 978; the nucleotide sequence of SEQ ID NO:13 from nucleotide 1084 to nucleotide 1854; the nucleotide sequence of the full-length protein coding sequence of clone fm60_1 deposited under accession number ATCC 98415; or the nucleotide sequence of a mature protein coding sequence of clone fm60_1 deposited under accession number ATCC 98415. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fm60_1 deposited under accession number ATCC 98415.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:13.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:14;

(b) fragments of the amino acid sequence of SEQ ID NO:14 comprising the amino acid sequence from amino acid 63 to amino acid 72 of SEQ ID NO:14; and

(c) the amino acid sequence encoded by the cDNA insert of clone fm60_1 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:14.

In one embodiment, the present invention provides a composition comprising an isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 16 to nucleotide 309;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 127 to nucleotide 309;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fr473_2 deposited under accession number ATCC 98415;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fr473_2 deposited under accession number ATCC 98415;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fr473_2 deposited under accession number ATCC 98415;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fr473_2 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising the amino acid sequence from amino acid 44 to amino acid 53 of SEQ ID NO:16;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

Preferably, such polynucleotide comprises the nucleotide sequence of SEQ ID NO:15 from nucleotide 16 to nucleotide 309; the nucleotide sequence of SEQ ID NO:15 from nucleotide 127 to nucleotide 309; the nucleotide sequence of the full-length protein coding sequence of clone fr473_2 deposited under accession number ATCC 98415; or the nucleotide sequence of a mature protein coding sequence of clone fr473_2 deposited under accession number ATCC 98415. In other preferred embodiments, the polynucleotide encodes the full-length or a mature protein encoded by the cDNA insert of clone fr473_2 deposited under accession number ATCC 98415. In yet other preferred embodiments, the present invention provides a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 58.

Other embodiments provide the gene corresponding to the cDNA sequence of SEQ ID NO:15.

In other embodiments, the present invention provides a composition comprising a protein, wherein said protein comprises an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 58;

(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising the amino acid sequence from amino acid 44 to amino acid 53 of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone fr473_2 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins. Preferably such protein comprises the amino acid sequence of SEQ ID NO:16 or the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 58.

In certain preferred embodiments, the polynucleotide is operably linked to an expression control sequence. The invention also provides a host cell, including bacterial, yeast, insect and mammalian cells, transformed with such polynucleotide compositions. Also provided by the present invention are organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein.

Processes are also provided for producing a protein, which comprise:

(a) growing a culture of the host cell transformed with such polynucleotide compositions in a suitable culture medium; and

(b) purifying the protein from the culture.

The protein produced according to such methods is also provided by the present invention. Preferred embodiments include those in which the protein produced by such process is a mature form of the protein.

Protein compositions of the present invention may further comprise a pharmaceutically acceptable carrier. Compositions comprising an antibody which specifically reacts with such protein are also provided by the present invention.

Methods are also provided for preventing, treating or ameliorating a medical condition which comprises administering to a mammalian subject a therapeutically

effective amount of a composition comprising a protein of the present invention and a pharmaceutically acceptable carrier.

BRIEF DESCRIPTION OF THE DRAWINGS

Figures 1A and 1B are schematic representations of the pED6 and pNOTs vectors, respectively, used for deposit of clones disclosed herein.

DETAILED DESCRIPTION

ISOLATED PROTEINS AND POLYNUCLEOTIDES

Nucleotide and amino acid sequences, as presently determined, are reported below for each clone and protein disclosed in the present application. The nucleotide sequence of each clone can readily be determined by sequencing of the deposited clone in accordance with known methods. The predicted amino acid sequence (both full-length and mature forms) can then be determined from such nucleotide sequence. The amino acid sequence of the protein encoded by a particular clone can also be determined by expression of the clone in a suitable host cell, collecting the protein and determining its sequence. For each disclosed protein applicants have identified what they have determined to be the reading frame best identifiable with sequence information available at the time of filing.

As used herein a "secreted" protein is one which, when expressed in a suitable host cell, is transported across or through a membrane, including transport as a result of signal sequences in its amino acid sequence. "Secreted" proteins include without limitation proteins secreted wholly (e.g., soluble proteins) or partially (e.g. , receptors) from the cell in which they are expressed. "Secreted" proteins also include without limitation proteins which are transported across the membrane of the endoplasmic reticulum.

Clone "ci25_4"

A polynucleotide of the present invention has been identified as clone "ci25_4". ci25_4 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. ci25_4 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "ci25_4 protein").

The nucleotide sequence of ci25_4 as presently determined is reported in SEQ ID NO:1. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the ci25_4 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:2. Amino acids 9 to 21 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 22, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone ci25_4 should be approximately 1700 bp.

The nucleotide sequence disclosed herein for ci25_4 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. ci25_4 demonstrated at least some similarity with sequences identified as AA243050 (zr24h03.r1 Stratagene NT2 neuronal precursor 937230 Homo sapiens cDNA clone 664373 5'), AA316800 (EST188485 HCC cell line (matastasis to liver in mouse) II Homo sapiens cDNA 5' end), AA340783 (EST46083 Fetal kidney II Homo sapiens cDNA 5' end), Q05686 (Islets of Langerhans cell clone ICA12.3 (ATCC 40703)), R12690 (yf40e07.s1 Homo sapiens cDNA clone 129348 3'), R16432 (yf40e07.r1 Homo sapiens cDNA clone), W81653 (zd84d12.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 347351 5'), and W81654 (zd84d12.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 347351 3'). Based upon sequence similarity, ci25_4 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts five additional potential transmembrane domains within the ci25_4 protein sequence, centered around amino acids 81, 134, 159, 182, and 241 of SEQ ID NO:2, respectively.

Clone "da228_6"

A polynucleotide of the present invention has been identified as clone "da228_6". da228_6 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. da228_6 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "da228_6 protein").

The nucleotide sequence of da228_6 as presently determined is reported in SEQ ID NO:3. What applicants presently believe to be the proper reading frame and the

predicted amino acid sequence of the da228_6 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:4.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone da228_6 should be approximately 1500 bp.

The nucleotide sequence disclosed herein for da228_6 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. da228_6 demonstrated at least some similarity with sequences identified as W57906 (zd17f11.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 340941 5') and W57907 (zd17f11.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 340941 3'. Based upon sequence similarity, da228_6 proteins and each similar protein or peptide may share at least some activity.

Clone "du410_5"

A polynucleotide of the present invention has been identified as clone "du410_5". du410_5 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. du410_5 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "du410_5 protein").

The nucleotide sequence of du410_5 as presently determined is reported in SEQ ID NO:5. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the du410_5 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:6.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone du410_5 should be approximately 2400 bp.

The nucleotide sequence disclosed herein for du410_5 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. du410_5 demonstrated at least some similarity with sequences identified as N44315 (EST51p19 WATM1 Homo sapiens cDNA clone 51p19) and N66980 (yz58d04.s1 Homo sapiens cDNA clone 287239 3'). The predicted amino acid sequence disclosed herein for du410_5 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted du410_5 protein demonstrated at least some similarity to sequences identified as U67604 (P115 protein

[Methanococcus jannaschii]). Based upon sequence similarity, du410_5 proteins and each similar protein or peptide may share at least some activity.

Clone "eh80_1"

A polynucleotide of the present invention has been identified as clone "eh80_1". eh80_1 was isolated from a human adult blood (peripheral blood mononuclear cells treated with granulocyte-colony stimulating factor *in vivo*) cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. eh80_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "eh80_1 protein").

The nucleotide sequence of eh80_1 as presently determined is reported in SEQ ID NO:7. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the eh80_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:8. Another potential eh80_1 reading frame and predicted amino acid sequence is encoded by basepairs 41 to 1659 of SEQ ID NO:7 and is reported in SEQ ID NO:25. A frameshift in the nucleotide sequence of SEQ ID NO:5 between about nucleotide 41 to about nucleotide 614 could join together portions of the overlapping reading frames of SEQ ID NO:8 and SEQ ID NO:25.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone eh80_1 should be approximately 2000 bp.

The nucleotide sequence disclosed herein for eh80_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. eh80_1 demonstrated at least some similarity with sequences identified as AA012957 (ze27b03.r1 Soares retina N2b4HR Homo sapiens cDNA clone 360173 5'), AA019878 (ze63b03.s1 Soares retina N2b4HR Homo sapiens cDNA clone 363629 3'), AA505456 (nh84c07.s1 NCI_CGAP_Br1.1 Homo sapiens cDNA clone IMAGE 965196), Q60246 (Human brain Expressed Sequence Tag EST02242), R16603 (yf43c04.r1 Homo sapiens cDNA clone 129606 5'), and T85469 (yd82f05.r1 Homo sapiens cDNA clone 114753 5'). The predicted amino acid sequence disclosed herein for eh80_1 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted eh80_1 protein demonstrated at least some similarity to sequences identified as U40747 (FBP 11 [Mus musculus]). Based upon sequence

similarity, eh80_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts two potential transmembrane domains within the amino acid sequence of SEQ ID NO:8, one centered around amino acid 107 and another around amino acid 131.

Clone "er369_1"

A polynucleotide of the present invention has been identified as clone "er369_1". er369_1 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. er369_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "er369_1 protein").

The nucleotide sequence of er369_1 as presently determined is reported in SEQ ID NO:9. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the er369_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:10. Amino acids 17 to 29 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 30, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone er369_1 should be approximately 1500 bp.

The nucleotide sequence disclosed herein for er369_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. er369_1 demonstrated at least some similarity with sequences identified as H12227 (ym12g10.r1 Homo sapiens cDNA clone 47729 5'), H70978 (yr73g06.r1 Homo sapiens cDNA clone 210970 5'), M79179 (EST01327 Homo sapiens cDNA clone HHCPO81), Q61324 (Human brain Expressed Sequence Tag EST01327), and R53554 (yg84e04.s1 Homo sapiens cDNA clone 39854 3' similar to contains Alu repetitive element). Based upon sequence similarity, er369_1 proteins and each similar protein or peptide may share at least some activity. The nucleotide sequence of er369_1 indicates that it may contain an Alu repetitive element.

Clone "fh123_5"

A polynucleotide of the present invention has been identified as clone "fh123_5". fh123_5 was isolated from a human fetal brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. fh123_5 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fh123_5 protein").

The nucleotide sequence of fh123_5 as presently determined is reported in SEQ ID NO:11. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the fh123_5 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:12. Amino acids 694 to 706 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 707, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fh123_5 should be approximately 2800 bp.

The nucleotide sequence disclosed herein for fh123_5 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fh123_5 demonstrated at least some similarity with sequences identified as AA815253 (ai64d02.s1 Soares testis NHT Homo sapiens cDNA clone 1375587 3'), AA855689 (vw71h04.r1 Stratagene mouse heart (#937316) Mus musculus cDNA clone 1260439 5'), and W80785 (zd83d07.s1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 347245 3). The predicted amino acid sequence disclosed herein for fh123_5 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted fh123_5 protein demonstrated at least some similarity to sequences identified as D80005 (KIAA0183 [Homo sapiens]). Based upon sequence similarity, fh123_5 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts five additional possible transmembrane domains within the fh123_5 protein sequence.

Clone "fm60_1"

A polynucleotide of the present invention has been identified as clone "fm60_1". fm60_1 was isolated from a human adult brain cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was

identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. fm60_1 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fm60_1 protein").

The nucleotide sequence of fm60_1 as presently determined is reported in SEQ ID NO:13. What applicants presently believe to be the proper reading frame and the predicted amino acid sequence of the fm60_1 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:14.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fm60_1 should be approximately 2200 bp.

The nucleotide sequence disclosed herein for fm60_1 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fm60_1 demonstrated at least some similarity with sequences identified as AA155574 (zo70a01.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 592200 3'), AF015147 (Homo sapiens clone HS19.1 Alu-Ya5 sequence), N86095 (J6377F Fetal heart, Lambda ZAP Express Homo sapiens cDNA clone J6377 5' similar to REPETITIVE ELEMENT ALU), U14567 (***ALU WARNING Human Alu-J subfamily consensus sequence), and Z82199 (Human DNA sequence from clone J316D5). Based upon sequence similarity, fm60_1 proteins and each similar protein or peptide may share at least some activity. The TopPredII computer program predicts a potential transmembrane domain within the fm60_1 protein sequence centered around amino acid 50 of SEQ ID NO:14. The nucleotide sequence of fm60_1 indicates that it may contain one or more of the following repetitive elements: Alu, L1.

Clone "fr473_2"

A polynucleotide of the present invention has been identified as clone "fr473_2". fr473_2 was isolated from a human adult placenta cDNA library using methods which are selective for cDNAs encoding secreted proteins (see U.S. Pat. No. 5,536,637), or was identified as encoding a secreted or transmembrane protein on the basis of computer analysis of the amino acid sequence of the encoded protein. fr473_2 is a full-length clone, including the entire coding sequence of a secreted protein (also referred to herein as "fr473_2 protein").

The nucleotide sequence of fr473_2 as presently determined is reported in SEQ ID NO:15. What applicants presently believe to be the proper reading frame and the

predicted amino acid sequence of the fr473_2 protein corresponding to the foregoing nucleotide sequence is reported in SEQ ID NO:16. Amino acids 25 to 37 are a predicted leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 38, or are a transmembrane domain. Amino acids 62 to 74 are another possible leader/signal sequence, with the predicted mature amino acid sequence beginning at amino acid 75, or are a transmembrane domain.

The EcoRI/NotI restriction fragment obtainable from the deposit containing clone fr473_2 should be approximately 605 bp.

The nucleotide sequence disclosed herein for fr473_2 was searched against the GenBank and GeneSeq nucleotide sequence databases using BLASTN/BLASTX and FASTA search protocols. fr473_2 demonstrated at least some similarity with sequences identified as AA479559 (zu42a02.r1 Soares ovary tumor NbHOT Homo sapiens cDNA clone 740618 5' similar to WP:F49C12.12 CE03372), H46855 (yo18g04.r1 Homo sapiens cDNA clone 178326 5'), T24372 (Human gene signature HUMGS06404), W31692 (zb93d01.r1 Soares parathyroid tumor NbHPA Homo sapiens cDNA clone 320353 5'), and Z32877 (H. sapiens partial cDNA sequence; clone HEA41P; single read). The predicted amino acid sequence disclosed herein for fr473_2 was searched against the GenPept and GeneSeq amino acid sequence databases using the BLASTX search protocol. The predicted fr473_2 protein demonstrated at least some similarity to sequences identified as Z68227 (F49C12.12 [Caenorhabditis elegans]). Based upon sequence similarity, fr473_2 proteins and each similar protein or peptide may share at least some activity.

Deposit of Clones

Clones ci25_4, da228_6, du410_5, eh80_1, er369_1, fh123_5, fm60_1, and fr473_2 were deposited on April 25, 1997 with the American Type Culture Collection (10801 University Boulevard, Manassas, Virginia 20110-2209 U.S.A.) as an original deposit under the Budapest Treaty and were given the accession number ATCC 98415, from which each clone comprising a particular polynucleotide is obtainable. All restrictions on the availability to the public of the deposited material will be irrevocably removed upon the granting of the patent, except for the requirements specified in 37 C.F.R. § 1.808(b), and the term of the deposit will comply with 37 C.F.R. § 1.806.

Each clone has been transfected into separate bacterial cells (*E. coli*) in this composite deposit. Each clone can be removed from the vector in which it was deposited by performing an EcoRI/NotI digestion (5' site, EcoRI; 3' site, NotI) to produce the

appropriate fragment for such clone. Each clone was deposited in either the pED6 or pNOTs vector depicted in Figures 1A and 1B, respectively. The pED6dpc2 vector ("pED6") was derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning (Kaufman *et al.*, 1991, *Nucleic Acids Res.* 19: 4485-4490); the pNOTs vector was derived from pMT2 (Kaufman *et al.*, 1989, *Mol. Cell. Biol.* 9: 946-958) by deletion of the DHFR sequences, insertion of a new polylinker, and insertion of the M13 origin of replication in the ClaI site. In some instances, the deposited clone can become "flipped" (i.e., in the reverse orientation) in the deposited isolate. In such instances, the cDNA insert can still be isolated by digestion with EcoRI and NotI. However, NotI will then produce the 5' site and EcoRI will produce the 3' site for placement of the cDNA in proper orientation for expression in a suitable vector. The cDNA may also be expressed from the vectors in which they were deposited.

Bacterial cells containing a particular clone can be obtained from the composite deposit as follows:

An oligonucleotide probe or probes should be designed to the sequence that is known for that particular clone. This sequence can be derived from the sequences provided herein, or from a combination of those sequences. The sequence of an oligonucleotide probe that was used to isolate or to sequence each full-length clone is identified below, and should be most reliable in isolating the clone of interest.

Clone	Probe Sequence
ci25_4	SEQ ID NO:17
da228_6	SEQ ID NO:18
du410_5	SEQ ID NO:19
eh80_1	SEQ ID NO:20
er369_1	SEQ ID NO:21
fh123_5	SEQ ID NO:22
fm60_1	SEQ ID NO:23
fr473_2	SEQ ID NO:24

In the sequences listed above which include an N at position 2, that position is occupied in preferred probes/primers by a biotinylated phosphoaramidite residue rather than a nucleotide (such as , for example, that produced by use of biotin phosphoramidite (1-

dimethoxytrityloxy-2-(N-biotinyl-4-aminobutyl)-propyl-3-O-(2-cyanoethyl)-(N,N-diisopropyl)-phosphoramadite) (Glen Research, cat. no. 10-1953)).

The design of the oligonucleotide probe should preferably follow these parameters:

(a) It should be designed to an area of the sequence which has the fewest ambiguous bases ("N's"), if any;

(b) It should be designed to have a T_m of approx. 80 ° C (assuming 2° for each A or T and 4 degrees for each G or C).

The oligonucleotide should preferably be labeled with g-^{32}P ATP (specific activity 6000 Ci/mmole) and T4 polynucleotide kinase using commonly employed techniques for labeling oligonucleotides. Other labeling techniques can also be used. Unincorporated label should preferably be removed by gel filtration chromatography or other established methods. The amount of radioactivity incorporated into the probe should be quantitated by measurement in a scintillation counter. Preferably, specific activity of the resulting probe should be approximately 4e+6 dpm/pmole.

The bacterial culture containing the pool of full-length clones should preferably be thawed and 100 µl of the stock used to inoculate a sterile culture flask containing 25 ml of sterile L-broth containing ampicillin at 100 µg/ml. The culture should preferably be grown to saturation at 37°C, and the saturated culture should preferably be diluted in fresh L-broth. Aliquots of these dilutions should preferably be plated to determine the dilution and volume which will yield approximately 5000 distinct and well-separated colonies on solid bacteriological media containing L-broth containing ampicillin at 100 µg/ml and agar at 1.5% in a 150 mm petri dish when grown overnight at 37°C. Other known methods of obtaining distinct, well-separated colonies can also be employed.

Standard colony hybridization procedures should then be used to transfer the colonies to nitrocellulose filters and lyse, denature and bake them.

The filter is then preferably incubated at 65°C for 1 hour with gentle agitation in 6X SSC (20X stock is 175.3 g NaCl/liter, 88.2 g Na citrate/liter, adjusted to pH 7.0 with NaOH) containing 0.5% SDS, 100 µg/ml of yeast RNA, and 10 mM EDTA (approximately 10 mL per 150 mm filter). Preferably, the probe is then added to the hybridization mix at a concentration greater than or equal to 1e+6 dpm/mL. The filter is then preferably incubated at 65°C with gentle agitation overnight. The filter is then preferably washed in 500 mL of 2X SSC/0.5% SDS at room temperature without agitation, preferably followed by 500 mL of 2X SSC/0.1% SDS at room temperature with gentle shaking for 15 minutes.

A third wash with 0.1X SSC/0.5% SDS at 65°C for 30 minutes to 1 hour is optional. The filter is then preferably dried and subjected to autoradiography for sufficient time to visualize the positives on the X-ray film. Other known hybridization methods can also be employed.

The positive colonies are picked, grown in culture, and plasmid DNA isolated using standard procedures. The clones can then be verified by restriction analysis, hybridization analysis, or DNA sequencing.

Fragments of the proteins of the present invention which are capable of exhibiting biological activity are also encompassed by the present invention. Fragments of the protein may be in linear form or they may be cyclized using known methods, for example, as described in H.U. Saragovi, *et al.*, Bio/Technology 10, 773-778 (1992) and in R.S. McDowell, *et al.*, J. Amer. Chem. Soc. 114, 9245-9253 (1992), both of which are incorporated herein by reference. Such fragments may be fused to carrier molecules such as immunoglobulins for many purposes, including increasing the valency of protein binding sites. For example, fragments of the protein may be fused through "linker" sequences to the Fc portion of an immunoglobulin. For a bivalent form of the protein, such a fusion could be to the Fc portion of an IgG molecule. Other immunoglobulin isotypes may also be used to generate such fusions. For example, a protein - IgM fusion would generate a decavalent form of the protein of the invention.

The present invention also provides both full-length and mature forms of the disclosed proteins. The full-length form of the such proteins is identified in the sequence listing by translation of the nucleotide sequence of each disclosed clone. The mature form(s) of such protein may be obtained by expression of the disclosed full-length polynucleotide (preferably those deposited with ATCC) in a suitable mammalian cell or other host cell. The sequence(s) of the mature form(s) of the protein may also be determinable from the amino acid sequence of the full-length form.

The present invention also provides genes corresponding to the polynucleotide sequences disclosed herein. "Corresponding genes" are the regions of the genome that are transcribed to produce the mRNAs from which cDNA polynucleotide sequences are derived and may include contiguous regions of the genome necessary for the regulated expression of such genes. Corresponding genes may therefore include but are not limited to coding sequences, 5' and 3' untranslated regions, alternatively spliced exons, introns, promoters, enhancers, and silencer or suppressor elements. The corresponding genes can be isolated in accordance with known methods using the sequence information disclosed

herein. Such methods include the preparation of probes or primers from the disclosed sequence information for identification and/or amplification of genes in appropriate genomic libraries or other sources of genomic materials. An "isolated gene" is a gene that has been separated from the adjacent coding sequences, if any, present in the genome of the organism from which the gene was isolated.

Organisms that have enhanced, reduced, or modified expression of the gene(s) corresponding to the polynucleotide sequences disclosed herein are provided. The desired change in gene expression can be achieved through the use of antisense polynucleotides or ribozymes that bind and/or cleave the mRNA transcribed from the gene (Albert and Morris, 1994, *Trends Pharmacol. Sci.* **15(7)**: 250-254; Lavarosky *et al.*, 1997, *Biochem. Mol. Med.* **62(1)**: 11-22; and Hampel, 1998, *Prog. Nucleic Acid Res. Mol. Biol.* **58**: 1-39; all of which are incorporated by reference herein). Transgenic animals that have multiple copies of the gene(s) corresponding to the polynucleotide sequences disclosed herein, preferably produced by transformation of cells with genetic constructs that are stably maintained within the transformed cells and their progeny, are provided. Transgenic animals that have modified genetic control regions that increase or reduce gene expression levels, or that change temporal or spatial patterns of gene expression, are also provided (see European Patent No. 0 649 464 B1, incorporated by reference herein). In addition, organisms are provided in which the gene(s) corresponding to the polynucleotide sequences disclosed herein have been partially or completely inactivated, through insertion of extraneous sequences into the corresponding gene(s) or through deletion of all or part of the corresponding gene(s). Partial or complete gene inactivation can be accomplished through insertion, preferably followed by imprecise excision, of transposable elements (Plasterk, 1992, *Bioessays* **14(9)**: 629-633; Zwaal *et al.*, 1993, *Proc. Natl. Acad. Sci. USA* **90(16)**: 7431-7435; Clark *et al.*, 1994, *Proc. Natl. Acad. Sci. USA* **91(2)**: 719-722; all of which are incorporated by reference herein), or through homologous recombination, preferably detected by positive/negative genetic selection strategies (Mansour *et al.*, 1988, *Nature* **336**: 348-352; U.S. Patent Nos. 5,464,764; 5,487,992; 5,627,059; 5,631,153; 5,614, 396; 5,616,491; and 5,679,523; all of which are incorporated by reference herein). These organisms with altered gene expression are preferably eukaryotes and more preferably are mammals. Such organisms are useful for the development of non-human models for the study of disorders involving the corresponding gene(s), and for the development of assay systems for the identification of molecules that interact with the protein product(s) of the corresponding gene(s).

Where the protein of the present invention is membrane-bound (e.g., is a receptor), the present invention also provides for soluble forms of such protein. In such forms part or all of the intracellular and transmembrane domains of the protein are deleted such that the protein is fully secreted from the cell in which it is expressed. The intracellular and transmembrane domains of proteins of the invention can be identified in accordance with known techniques for determination of such domains from sequence information.

Proteins and protein fragments of the present invention include proteins with amino acid sequence lengths that are at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of a disclosed protein and have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with that disclosed protein, where sequence identity is determined by comparing the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Also included in the present invention are proteins and protein fragments that contain a segment preferably comprising 8 or more (more preferably 20 or more, most preferably 30 or more) contiguous amino acids that shares at least 75% sequence identity (more preferably, at least 85% identity; most preferably at least 95% identity) with any such segment of any of the disclosed proteins.

Species homologues of the disclosed polynucleotides and proteins are also provided by the present invention. As used herein, a "species homologue" is a protein or polynucleotide with a different species of origin from that of a given protein or polynucleotide, but with significant sequence similarity to the given protein or polynucleotide. Preferably, polynucleotide species homologues have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, and protein species homologues have at least 30% sequence identity (more preferably, at least 45% identity; most preferably at least 60% identity) with the given protein, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides or the amino acid sequences of the proteins when aligned so as to maximize overlap and identity while minimizing sequence gaps. Species homologues may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from the desired species. Preferably, species homologues are those isolated from mammalian species. Most preferably, species homologues are those isolated from certain mammalian species such as, for example, *Pan troglodytes, Gorilla gorilla, Pongo pygmaeus, Hylobates*

concolor, Macaca mulatta, Papio papio, Papio hamadryas, Cercopithecus aethiops, Cebus capucinus, Aotus trivirgatus, Sanguinus oedipus, Microcebus murinus, Mus musculus, Rattus norvegicus, Cricetulus griseus, Felis catus, Mustela vison, Canis familiaris, Oryctolagus cuniculus, Bos taurus, Ovis aries, Sus scrofa, and *Equus caballus,* for which genetic maps have been created allowing the identification of syntenic relationships between the genomic organization of genes in one species and the genomic organization of the related genes in another species (O'Brien and Seuánez, 1988, *Ann. Rev. Genet.* 22: 323-351; O'Brien *et al.*, 1993, *Nature Genetics* 3:103-112; Johansson *et al.*, 1995, *Genomics* 25: 682-690; Lyons *et al.*, 1997, *Nature Genetics* 15: 47-56; O'Brien *et al.*, 1997, *Trends in Genetics* 13(10): 393-399; Carver and Stubbs, 1997, *Genome Research* 7:1123-1137; all of which are incorporated by reference herein).

The invention also encompasses allelic variants of the disclosed polynucleotides or proteins; that is, naturally-occurring alternative forms of the isolated polynucleotides which also encode proteins which are identical or have significantly similar sequences to those encoded by the disclosed polynucleotides. Preferably, allelic variants have at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% identity) with the given polynucleotide, where sequence identity is determined by comparing the nucleotide sequences of the polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps. Allelic variants may be isolated and identified by making suitable probes or primers from the sequences provided herein and screening a suitable nucleic acid source from individuals of the appropriate species.

The invention also includes polynucleotides with sequences complementary to those of the polynucleotides disclosed herein.

The present invention also includes polynucleotides that hybridize under reduced stringency conditions, more preferably stringent conditions, and most preferably highly stringent conditions, to polynucleotides described herein. Examples of stringency conditions are shown in the table below: highly stringent conditions are those that are at least as stringent as, for example, conditions A-F; stringent conditions are at least as stringent as, for example, conditions G-L; and reduced stringency conditions are at least as stringent as, for example, conditions M-R.

Stringency Condition	Polynucleotide Hybrid	Hybrid Length (bp)‡	Hybridization Temperature and Buffer†	Wash Temperature and Buffer†
A	DNA:DNA	≥ 50	65°C; 1xSSC -or- 42°C; 1xSSC, 50% formamide	65°C; 0.3xSSC
B	DNA:DNA	<50	T_B*; 1xSSC	T_B*; 1xSSC
C	DNA:RNA	≥ 50	67°C; 1xSSC -or- 45°C; 1xSSC, 50% formamide	67°C; 0.3xSSC
D	DNA:RNA	<50	T_D*; 1xSSC	T_D*; 1xSSC
E	RNA:RNA	≥ 50	70°C; 1xSSC -or- 50°C; 1xSSC, 50% formamide	70°C; 0.3xSSC
F	RNA:RNA	<50	T_F*; 1xSSC	T_F*; 1xSSC
G	DNA:DNA	≥ 50	65°C; 4xSSC -or- 42°C; 4xSSC, 50% formamide	65°C; 1xSSC
H	DNA:DNA	<50	T_H*; 4xSSC	T_H*; 4xSSC
I	DNA:RNA	≥ 50	67°C; 4xSSC -or- 45°C; 4xSSC, 50% formamide	67°C; 1xSSC
J	DNA:RNA	<50	T_J*; 4xSSC	T_J*; 4xSSC
K	RNA:RNA	≥ 50	70°C; 4xSSC -or- 50°C; 4xSSC, 50% formamide	67°C; 1xSSC
L	RNA:RNA	<50	T_L*; 2xSSC	T_L*; 2xSSC
M	DNA:DNA	≥ 50	50°C; 4xSSC -or- 40°C; 6xSSC, 50% formamide	50°C; 2xSSC
N	DNA:DNA	<50	T_N*; 6xSSC	T_N*; 6xSSC
O	DNA:RNA	≥ 50	55°C; 4xSSC -or- 42°C; 6xSSC, 50% formamide	55°C; 2xSSC
P	DNA:RNA	<50	T_P*; 6xSSC	T_P*; 6xSSC
Q	RNA:RNA	≥ 50	60°C; 4xSSC -or- 45°C; 6xSSC, 50% formamide	60°C; 2xSSC
R	RNA:RNA	<50	T_R*; 4xSSC	T_R*; 4xSSC

‡: The hybrid length is that anticipated for the hybridized region(s) of the hybridizing polynucleotides. When hybridizing a polynucleotide to a target polynucleotide of unknown sequence, the hybrid length is assumed to be that of the hybridizing polynucleotide. When polynucleotides of known sequence are hybridized, the hybrid length can be determined by aligning the sequences of the polynucleotides and identifying the region or regions of optimal sequence complementarity.

†: SSPE (1xSSPE is 0.15M NaCl, 10mM NaH_2PO_4, and 1.25mM EDTA, pH 7.4) can be substituted for SSC (1xSSC is 0.15M NaCl and 15mM sodium citrate) in the hybridization and wash buffers; washes are performed for 15 minutes after hybridization is complete.

*T_B - T_R: The hybridization temperature for hybrids anticipated to be less than 50 base pairs in length should be 5-10°C less than the melting temperature (T_m) of the hybrid, where T_m is determined according to the following equations. For hybrids less than 18 base pairs in length, T_m(°C) = 2(# of A + T bases) + 4(# of G + C bases). For hybrids between 18 and 49 base pairs in length, T_m(°C) = 81.5 + 16.6($\log_{10}$[Na^+]) + 0.41(%G+C) - (600/N), where N is the number of bases in the hybrid, and [Na^+] is the concentration of sodium ions in the hybridization buffer ([Na^+] for 1xSSC = 0.165 M).

Additional examples of stringency conditions for polynucleotide hybridization are provided in Sambrook, J., E.F. Fritsch, and T. Maniatis, 1989, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, chapters 9 and 11, and *Current Protocols in Molecular Biology*, 1995, F.M. Ausubel et al., eds., John Wiley & Sons, Inc., sections 2.10 and 6.3-6.4, incorporated herein by reference.

Preferably, each such hybridizing polynucleotide has a length that is at least 25%(more preferably at least 50%, and most preferably at least 75%) of the length of the polynucleotide of the present invention to which it hybridizes, and has at least 60% sequence identity (more preferably, at least 75% identity; most preferably at least 90% or 95% identity) with the polynucleotide of the present invention to which it hybridizes, where sequence identity is determined by comparing the sequences of the hybridizing polynucleotides when aligned so as to maximize overlap and identity while minimizing sequence gaps.

The isolated polynucleotide of the invention may be operably linked to an expression control sequence such as the pMT2 or pED expression vectors disclosed in Kaufman *et al.*, Nucleic Acids Res. 19, 4485-4490 (1991), in order to produce the protein recombinantly. Many suitable expression control sequences are known in the art. General methods of expressing recombinant proteins are also known and are exemplified in R. Kaufman, Methods in Enzymology 185, 537-566 (1990). As defined herein "operably linked" means that the isolated polynucleotide of the invention and an expression control sequence are situated within a vector or cell in such a way that the protein is expressed by a host cell which has been transformed (transfected) with the ligated polynucleotide/expression control sequence.

A number of types of cells may act as suitable host cells for expression of the protein. Mammalian host cells include, for example, monkey COS cells, Chinese Hamster Ovary (CHO) cells, human kidney 293 cells, human epidermal A431 cells, human Colo205 cells, 3T3 cells, CV-1 cells, other transformed primate cell lines, normal diploid cells, cell strains derived from in vitro culture of primary tissue, primary explants, HeLa cells, mouse L cells, BHK, HL-60, U937, HaK or Jurkat cells.

Alternatively, it may be possible to produce the protein in lower eukaryotes such as yeast or in prokaryotes such as bacteria. Potentially suitable yeast strains include *Saccharomyces cerevisiae, Schizosaccharomyces pombe, Kluyveromyces* strains, *Candida*, or any yeast strain capable of expressing heterologous proteins. Potentially suitable bacterial strains include *Escherichia coli, Bacillus subtilis, Salmonella typhimurium*, or any bacterial

strain capable of expressing heterologous proteins. If the protein is made in yeast or bacteria, it may be necessary to modify the protein produced therein, for example by phosphorylation or glycosylation of the appropriate sites, in order to obtain the functional protein. Such covalent attachments may be accomplished using known chemical or enzymatic methods.

The protein may also be produced by operably linking the isolated polynucleotide of the invention to suitable control sequences in one or more insect expression vectors, and employing an insect expression system. Materials and methods for baculovirus/insect cell expression systems are commercially available in kit form from, *e.g.*, Invitrogen, San Diego, California, U.S.A. (the MaxBac® kit), and such methods are well known in the art, as described in Summers and Smith, Texas Agricultural Experiment Station Bulletin No. 1555 (1987), incorporated herein by reference. As used herein, an insect cell capable of expressing a polynucleotide of the present invention is "transformed."

The protein of the invention may be prepared by culturing transformed host cells under culture conditions suitable to express the recombinant protein. The resulting expressed protein may then be purified from such culture (i.e., from culture medium or cell extracts) using known purification processes, such as gel filtration and ion exchange chromatography. The purification of the protein may also include an affinity column containing agents which will bind to the protein; one or more column steps over such affinity resins as concanavalin A-agarose, heparin-toyopearl® or Cibacrom blue 3GA Sepharose®; one or more steps involving hydrophobic interaction chromatography using such resins as phenyl ether, butyl ether, or propyl ether; or immunoaffinity chromatography.

Alternatively, the protein of the invention may also be expressed in a form which will facilitate purification. For example, it may be expressed as a fusion protein, such as those of maltose binding protein (MBP), glutathione-S-transferase (GST) or thioredoxin (TRX). Kits for expression and purification of such fusion proteins are commercially available from New England BioLab (Beverly, MA), Pharmacia (Piscataway, NJ) and InVitrogen, respectively. The protein can also be tagged with an epitope and subsequently purified by using a specific antibody directed to such epitope. One such epitope ("Flag") is commercially available from Kodak (New Haven, CT).

Finally, one or more reverse-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media, e.g., silica gel having pendant

methyl or other aliphatic groups, can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, can also be employed to provide a substantially homogeneous isolated recombinant protein. The protein thus purified is substantially free of other mammalian proteins and is defined in accordance with the present invention as an "isolated protein."

The protein of the invention may also be expressed as a product of transgenic animals, e.g., as a component of the milk of transgenic cows, goats, pigs, or sheep which are characterized by somatic or germ cells containing a nucleotide sequence encoding the protein.

The protein may also be produced by known conventional chemical synthesis. Methods for constructing the proteins of the present invention by synthetic means are known to those skilled in the art. The synthetically-constructed protein sequences, by virtue of sharing primary, secondary or tertiary structural and/or conformational characteristics with proteins may possess biological properties in common therewith, including protein activity. Thus, they may be employed as biologically active or immunological substitutes for natural, purified proteins in screening of therapeutic compounds and in immunological processes for the development of antibodies.

The proteins provided herein also include proteins characterized by amino acid sequences similar to those of purified proteins but into which modification are naturally provided or deliberately engineered. For example, modifications in the peptide or DNA sequences can be made by those skilled in the art using known techniques. Modifications of interest in the protein sequences may include the alteration, substitution, replacement, insertion or deletion of a selected amino acid residue in the coding sequence. For example, one or more of the cysteine residues may be deleted or replaced with another amino acid to alter the conformation of the molecule. Techniques for such alteration, substitution, replacement, insertion or deletion are well known to those skilled in the art (see, e.g., U.S. Patent No. 4,518,584). Preferably, such alteration, substitution, replacement, insertion or deletion retains the desired activity of the protein.

Other fragments and derivatives of the sequences of proteins which would be expected to retain protein activity in whole or in part and may thus be useful for screening or other immunological methodologies may also be easily made by those skilled in the art given the disclosures herein. Such modifications are believed to be encompassed by the present invention.

USES AND BIOLOGICAL ACTIVITY

The polynucleotides and proteins of the present invention are expected to exhibit one or more of the uses or biological activities (including those associated with assays cited herein) identified below. Uses or activities described for proteins of the present invention may be provided by administration or use of such proteins or by administration or use of polynucleotides encoding such proteins (such as, for example, in gene therapies or vectors suitable for introduction of DNA).

Research Uses and Utilities

The polynucleotides provided by the present invention can be used by the research community for various purposes. The polynucleotides can be used to express recombinant protein for analysis, characterization or therapeutic use; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in disease states); as molecular weight markers on Southern gels; as chromosome markers or tags (when labeled) to identify chromosomes or to map related gene positions; to compare with endogenous DNA sequences in patients to identify potential genetic disorders; as probes to hybridize and thus discover novel, related DNA sequences; as a source of information to derive PCR primers for genetic fingerprinting; as a probe to "subtract-out" known sequences in the process of discovering other novel polynucleotides; for selecting and making oligomers for attachment to a "gene chip" or other support, including for examination of expression patterns; to raise anti-protein antibodies using DNA immunization techniques; and as an antigen to raise anti-DNA antibodies or elicit another immune response. Where the polynucleotide encodes a protein which binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the polynucleotide can also be used in interaction trap assays (such as, for example, those described in Gyuris *et al.*, 1993, *Cell* 75: 791-803 and in Rossi *et al.*, 1997, *Proc. Natl. Acad. Sci. USA* 94: 8405-8410, all of which are incorporated by reference herein) to identify polynucleotides encoding the other protein with which binding occurs or to identify inhibitors of the binding interaction.

The proteins provided by the present invention can similarly be used in assay to determine biological activity, including in a panel of multiple proteins for high-throughput screening; to raise antibodies or to elicit another immune response; as a reagent (including the labeled reagent) in assays designed to quantitatively determine

levels of the protein (or its receptor) in biological fluids; as markers for tissues in which the corresponding protein is preferentially expressed (either constitutively or at a particular stage of tissue differentiation or development or in a disease state); and, of course, to isolate correlative receptors or ligands. Where the protein binds or potentially binds to another protein (such as, for example, in a receptor-ligand interaction), the protein can be used to identify the other protein with which binding occurs or to identify inhibitors of the binding interaction. Proteins involved in these binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction.

Any or all of these research utilities are capable of being developed into reagent grade or kit format for commercialization as research products.

Methods for performing the uses listed above are well known to those skilled in the art. References disclosing such methods include without limitation "Molecular Cloning: A Laboratory Manual", 2d ed., Cold Spring Harbor Laboratory Press, Sambrook, J., E.F. Fritsch and T. Maniatis eds., 1989, and "Methods in Enzymology: Guide to Molecular Cloning Techniques", Academic Press, Berger, S.L. and A.R. Kimmel eds., 1987.

Nutritional Uses

Polynucleotides and proteins of the present invention can also be used as nutritional sources or supplements. Such uses include without limitation use as a protein or amino acid supplement, use as a carbon source, use as a nitrogen source and use as a source of carbohydrate. In such cases the protein or polynucleotide of the invention can be added to the feed of a particular organism or can be administered as a separate solid or liquid preparation, such as in the form of powder, pills, solutions, suspensions or capsules. In the case of microorganisms, the protein or polynucleotide of the invention can be added to the medium in or on which the microorganism is cultured.

Cytokine and Cell Proliferation/Differentiation Activity

A protein of the present invention may exhibit cytokine, cell proliferation (either inducing or inhibiting) or cell differentiation (either inducing or inhibiting) activity or may induce production of other cytokines in certain cell populations. Many protein factors discovered to date, including all known cytokines, have exhibited activity in one or more factor dependent cell proliferation assays, and hence the assays serve as a convenient confirmation of cytokine activity. The activity of a protein of the present invention is

evidenced by any one of a number of routine factor dependent cell proliferation assays for cell lines including, without limitation, 32D, DA2, DA1G, T10, B9, B9/11, BaF3, MC9/G, M+ (preB M+), 2E8, RB5, DA1, 123, T1165, HT2, CTLL2, TF-1, Mo7e and CMK.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for T-cell or thymocyte proliferation include without limitation those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Bertagnolli et al., J. Immunol. 145:1706-1712, 1990; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Bertagnolli, et al., J. Immunol. 149:3778-3783, 1992; Bowman et al., J. Immunol. 152: 1756-1761, 1994.

Assays for cytokine production and/or proliferation of spleen cells, lymph node cells or thymocytes include, without limitation, those described in: Polyclonal T cell stimulation, Kruisbeek, A.M. and Shevach, E.M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.12.1-3.12.14, John Wiley and Sons, Toronto. 1994; and Measurement of mouse and human Interferon γ, Schreiber, R.D. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.8.1-6.8.8, John Wiley and Sons, Toronto. 1994.

Assays for proliferation and differentiation of hematopoietic and lymphopoietic cells include, without limitation, those described in: Measurement of Human and Murine Interleukin 2 and Interleukin 4, Bottomly, K., Davis, L.S. and Lipsky, P.E. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.3.1-6.3.12, John Wiley and Sons, Toronto. 1991; deVries et al., J. Exp. Med. 173:1205-1211, 1991; Moreau et al., Nature 336:690-692, 1988; Greenberger et al., Proc. Natl. Acad. Sci. U.S.A. 80:2931-2938, 1983; Measurement of mouse and human interleukin 6 - Nordan, R. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.6.1-6.6.5, John Wiley and Sons, Toronto. 1991; Smith et al., Proc. Natl. Acad. Sci. U.S.A. 83:1857-1861, 1986; Measurement of human Interleukin 11 - Bennett, F., Giannotti, J., Clark, S.C. and Turner, K. J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.15.1 John Wiley and Sons, Toronto. 1991; Measurement of mouse and human Interleukin 9 - Ciarletta, A., Giannotti, J., Clark, S.C. and Turner, K.J. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 6.13.1, John Wiley and Sons, Toronto. 1991.

Assays for T-cell clone responses to antigens (which will identify, among others, proteins that affect APC-T cell interactions as well as direct T-cell effects by measuring proliferation and cytokine production) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function; Chapter 6, Cytokines and their cellular receptors; Chapter 7, Immunologic studies in Humans); Weinberger et al., Proc. Natl. Acad. Sci. USA 77:6091-6095, 1980; Weinberger et al., Eur. J. Immun. 11:405-411, 1981; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988.

Immune Stimulating or Suppressing Activity

A protein of the present invention may also exhibit immune stimulating or immune suppressing activity, including without limitation the activities for which assays are described herein. A protein may be useful in the treatment of various immune deficiencies and disorders (including severe combined immunodeficiency (SCID)), e.g., in regulating (up or down) growth and proliferation of T and/or B lymphocytes, as well as effecting the cytolytic activity of NK cells and other cell populations. These immune deficiencies may be genetic or be caused by viral (e.g., HIV) as well as bacterial or fungal infections, or may result from autoimmune disorders. More specifically, infectious diseases causes by viral, bacterial, fungal or other infection may be treatable using a protein of the present invention, including infections by HIV, hepatitis viruses, herpesviruses, mycobacteria, Leishmania spp., malaria spp. and various fungal infections such as candidiasis. Of course, in this regard, a protein of the present invention may also be useful where a boost to the immune system generally may be desirable, *i.e.*, in the treatment of cancer.

Autoimmune disorders which may be treated using a protein of the present invention include, for example, connective tissue disease, multiple sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune pulmonary inflammation, Guillain-Barre syndrome, autoimmune thyroiditis, insulin dependent diabetes mellitis, myasthenia gravis, graft-versus-host disease and autoimmune inflammatory eye disease. Such a protein of the present invention may also to be useful in the treatment of allergic reactions and conditions, such as asthma (particularly allergic asthma) or other respiratory problems. Other conditions, in which immune suppression is desired (including, for

example, organ transplantation), may also be treatable using a protein of the present invention.

Using the proteins of the invention it may also be possible to immune responses, in a number of ways. Down regulation may be in the form of inhibiting or blocking an immune response already in progress or may involve preventing the induction of an immune response. The functions of activated T cells may be inhibited by suppressing T cell responses or by inducing specific tolerance in T cells, or both. Immunosuppression of T cell responses is generally an active, non-antigen-specific, process which requires continuous exposure of the T cells to the suppressive agent. Tolerance, which involves inducing non-responsiveness or anergy in T cells, is distinguishable from immunosuppression in that it is generally antigen-specific and persists after exposure to the tolerizing agent has ceased. Operationally, tolerance can be demonstrated by the lack of a T cell response upon reexposure to specific antigen in the absence of the tolerizing agent.

Down regulating or preventing one or more antigen functions (including without limitation B lymphocyte antigen functions (such as , for example, B7)), *e.g.*, preventing high level lymphokine synthesis by activated T cells, will be useful in situations of tissue, skin and organ transplantation and in graft-versus-host disease (GVHD). For example, blockage of T cell function should result in reduced tissue destruction in tissue transplantation. Typically, in tissue transplants, rejection of the transplant is initiated through its recognition as foreign by T cells, followed by an immune reaction that destroys the transplant. The administration of a molecule which inhibits or blocks interaction of a B7 lymphocyte antigen with its natural ligand(s) on immune cells (such as a soluble, monomeric form of a peptide having B7-2 activity alone or in conjunction with a monomeric form of a peptide having an activity of another B lymphocyte antigen (*e.g.*, B7-1, B7-3) or blocking antibody), prior to transplantation can lead to the binding of the molecule to the natural ligand(s) on the immune cells without transmitting the corresponding costimulatory signal. Blocking B lymphocyte antigen function in this matter prevents cytokine synthesis by immune cells, such as T cells, and thus acts as an immunosuppressant. Moreover, the lack of costimulation may also be sufficient to anergize the T cells, thereby inducing tolerance in a subject. Induction of long-term tolerance by B lymphocyte antigen-blocking reagents may avoid the necessity of repeated administration of these blocking reagents. To achieve sufficient immunosuppression or

tolerance in a subject, it may also be necessary to block the function of a combination of B lymphocyte antigens.

The efficacy of particular blocking reagents in preventing organ transplant rejection or GVHD can be assessed using animal models that are predictive of efficacy in humans. Examples of appropriate systems which can be used include allogeneic cardiac grafts in rats and xenogeneic pancreatic islet cell grafts in mice, both of which have been used to examine the immunosuppressive effects of CTLA4Ig fusion proteins *in vivo* as described in Lenschow *et al.*, Science 257:789-792 (1992) and Turka *et al.*, Proc. Natl. Acad. Sci USA, *89*:11102-11105 (1992). In addition, murine models of GVHD (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 846-847) can be used to determine the effect of blocking B lymphocyte antigen function *in vivo* on the development of that disease.

Blocking antigen function may also be therapeutically useful for treating autoimmune diseases. Many autoimmune disorders are the result of inappropriate activation of T cells that are reactive against self tissue and which promote the production of cytokines and autoantibodies involved in the pathology of the diseases. Preventing the activation of autoreactive T cells may reduce or eliminate disease symptoms. Administration of reagents which block costimulation of T cells by disrupting receptor:ligand interactions of B lymphocyte antigens can be used to inhibit T cell activation and prevent production of autoantibodies or T cell-derived cytokines which may be involved in the disease process. Additionally, blocking reagents may induce antigen-specific tolerance of autoreactive T cells which could lead to long-term relief from the disease. The efficacy of blocking reagents in preventing or alleviating autoimmune disorders can be determined using a number of well-characterized animal models of human autoimmune diseases. Examples include murine experimental autoimmune encephalitis, systemic lupus erythmatosis in MRL/*lpr/lpr* mice or NZB hybrid mice, murine autoimmune collagen arthritis, diabetes mellitus in NOD mice and BB rats, and murine experimental myasthenia gravis (see Paul ed., Fundamental Immunology, Raven Press, New York, 1989, pp. 840-856).

Upregulation of an antigen function (preferably a B lymphocyte antigen function), as a means of up regulating immune responses, may also be useful in therapy. Upregulation of immune responses may be in the form of enhancing an existing immune response or eliciting an initial immune response. For example, enhancing an immune response through stimulating B lymphocyte antigen function may be useful in cases of

viral infection. In addition, systemic viral diseases such as influenza, the common cold, and encephalitis might be alleviated by the administration of stimulatory forms of B lymphocyte antigens systemically.

Alternatively, anti-viral immune responses may be enhanced in an infected patient by removing T cells from the patient, costimulating the T cells *in vitro* with viral antigen-pulsed APCs either expressing a peptide of the present invention or together with a stimulatory form of a soluble peptide of the present invention and reintroducing the *in vitro* activated T cells into the patient. Another method of enhancing anti-viral immune responses would be to isolate infected cells from a patient, transfect them with a nucleic acid encoding a protein of the present invention as described herein such that the cells express all or a portion of the protein on their surface, and reintroduce the transfected cells into the patient. The infected cells would now be capable of delivering a costimulatory signal to, and thereby activate, T cells *in vivo*.

In another application, up regulation or enhancement of antigen function (preferably B lymphocyte antigen function) may be useful in the induction of tumor immunity. Tumor cells (*e.g.*, sarcoma, melanoma, lymphoma, leukemia, neuroblastoma, carcinoma) transfected with a nucleic acid encoding at least one peptide of the present invention can be administered to a subject to overcome tumor-specific tolerance in the subject. If desired, the tumor cell can be transfected to express a combination of peptides. For example, tumor cells obtained from a patient can be transfected *ex vivo* with an expression vector directing the expression of a peptide having B7-2-like activity alone, or in conjunction with a peptide having B7-1-like activity and/or B7-3-like activity. The transfected tumor cells are returned to the patient to result in expression of the peptides on the surface of the transfected cell. Alternatively, gene therapy techniques can be used to target a tumor cell for transfection *in vivo*.

The presence of the peptide of the present invention having the activity of a B lymphocyte antigen(s) on the surface of the tumor cell provides the necessary costimulation signal to T cells to induce a T cell mediated immune response against the transfected tumor cells. In addition, tumor cells which lack MHC class I or MHC class II molecules, or which fail to reexpress sufficient amounts of MHC class I or MHC class II molecules, can be transfected with nucleic acid encoding all or a portion of (*e.g.*, a cytoplasmic-domain truncated portion) of an MHC class I α chain protein and β_2 microglobulin protein or an MHC class II α chain protein and an MHC class II β chain protein to thereby express MHC class I or MHC class II proteins on the cell surface.

Expression of the appropriate class I or class II MHC in conjunction with a peptide having the activity of a B lymphocyte antigen (*e.g.*, B7-1, B7-2, B7-3) induces a T cell mediated immune response against the transfected tumor cell. Optionally, a gene encoding an antisense construct which blocks expression of an MHC class II associated protein, such as the invariant chain, can also be cotransfected with a DNA encoding a peptide having the activity of a B lymphocyte antigen to promote presentation of tumor associated antigens and induce tumor specific immunity. Thus, the induction of a T cell mediated immune response in a human subject may be sufficient to overcome tumor-specific tolerance in the subject.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for thymocyte or splenocyte cytotoxicity include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter 7, Immunologic studies in Humans); Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Herrmann et al., Proc. Natl. Acad. Sci. USA 78:2488-2492, 1981; Herrmann et al., J. Immunol. 128:1968-1974, 1982; Handa et al., J. Immunol. 135:1564-1572, 1985; Takai et al., J. Immunol. 137:3494-3500, 1986; Bowmanet al., J. Virology 61:1992-1998; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., Cellular Immunology 133:327-341, 1991; Brown et al., J. Immunol. 153:3079-3092, 1994.

Assays for T-cell-dependent immunoglobulin responses and isotype switching (which will identify, among others, proteins that modulate T-cell dependent antibody responses and that affect Th1/Th2 profiles) include, without limitation, those described in: Maliszewski, J. Immunol. 144:3028-3033, 1990; and Assays for B cell function: *In vitro* antibody production, Mond, J.J. and Brunswick, M. In *Current Protocols in Immunology*. J.E.e.a. Coligan eds. Vol 1 pp. 3.8.1-3.8.16, John Wiley and Sons, Toronto. 1994.

Mixed lymphocyte reaction (MLR) assays (which will identify, among others, proteins that generate predominantly Th1 and CTL responses) include, without limitation, those described in: Current Protocols in Immunology, Ed by J. E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 3, In Vitro assays for Mouse Lymphocyte Function 3.1-3.19; Chapter

7, Immunologic studies in Humans); Takai et al., J. Immunol. 137:3494-3500, 1986; Takai et al., J. Immunol. 140:508-512, 1988; Bertagnolli et al., J. Immunol. 149:3778-3783, 1992.

Dendritic cell-dependent assays (which will identify, among others, proteins expressed by dendritic cells that activate naive T-cells) include, without limitation, those described in: Guery et al., J. Immunol. 134:536-544, 1995; Inaba et al., Journal of Experimental Medicine 173:549-559, 1991; Macatonia et al., Journal of Immunology 154:5071-5079, 1995; Porgador et al., Journal of Experimental Medicine 182:255-260, 1995; Nair et al., Journal of Virology 67:4062-4069, 1993; Huang et al., Science 264:961-965, 1994; Macatonia et al., Journal of Experimental Medicine 169:1255-1264, 1989; Bhardwaj et al., Journal of Clinical Investigation 94:797-807, 1994; and Inaba et al., Journal of Experimental Medicine 172:631-640, 1990.

Assays for lymphocyte survival/apoptosis (which will identify, among others, proteins that prevent apoptosis after superantigen induction and proteins that regulate lymphocyte homeostasis) include, without limitation, those described in: Darzynkiewicz et al., Cytometry 13:795-808, 1992; Gorczyca et al., Leukemia 7:659-670, 1993; Gorczyca et al., Cancer Research 53:1945-1951, 1993; Itoh et al., Cell 66:233-243, 1991; Zacharchuk, Journal of Immunology 145:4037-4045, 1990; Zamai et al., Cytometry 14:891-897, 1993; Gorczyca et al., International Journal of Oncology 1:639-648, 1992.

Assays for proteins that influence early steps of T-cell commitment and development include, without limitation, those described in: Antica et al., Blood 84:111-117, 1994; Fine et al., Cellular Immunology 155:111-122, 1994; Galy et al., Blood 85:2770-2778, 1995; Toki et al., Proc. Nat. Acad Sci. USA 88:7548-7551, 1991.

Hematopoiesis Regulating Activity

A protein of the present invention may be useful in regulation of hematopoiesis and, consequently, in the treatment of myeloid or lymphoid cell deficiencies. Even marginal biological activity in support of colony forming cells or of factor-dependent cell lines indicates involvement in regulating hematopoiesis, e.g. in supporting the growth and proliferation of erythroid progenitor cells alone or in combination with other cytokines, thereby indicating utility, for example, in treating various anemias or for use in conjunction with irradiation/chemotherapy to stimulate the production of erythroid precursors and/or erythroid cells; in supporting the growth and proliferation of myeloid cells such as granulocytes and monocytes/macrophages (i.e., traditional CSF activity) useful, for example, in conjunction with chemotherapy to prevent or treat consequent

myelo-suppression; in supporting the growth and proliferation of megakaryocytes and consequently of platelets thereby allowing prevention or treatment of various platelet disorders such as thrombocytopenia, and generally for use in place of or complimentary to platelet transfusions; and/or in supporting the growth and proliferation of hematopoietic stem cells which are capable of maturing to any and all of the above-mentioned hematopoietic cells and therefore find therapeutic utility in various stem cell disorders (such as those usually treated with transplantation, including, without limitation, aplastic anemia and paroxysmal nocturnal hemoglobinuria), as well as in repopulating the stem cell compartment post irradiation/chemotherapy, either *in-vivo* or *ex-vivo* (i.e., in conjunction with bone marrow transplantation or with peripheral progenitor cell transplantation (homologous or heterologous)) as normal cells or genetically manipulated for gene therapy.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for proliferation and differentiation of various hematopoietic lines are cited above.

Assays for embryonic stem cell differentiation (which will identify, among others, proteins that influence embryonic differentiation hematopoiesis) include, without limitation, those described in: Johansson et al. Cellular Biology 15:141-151, 1995; Keller et al., Molecular and Cellular Biology 13:473-486, 1993; McClanahan et al., Blood 81:2903-2915, 1993.

Assays for stem cell survival and differentiation (which will identify, among others, proteins that regulate lympho-hematopoiesis) include, without limitation, those described in: Methylcellulose colony forming assays, Freshney, M.G. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 265-268, Wiley-Liss, Inc., New York, NY. 1994; Hirayama et al., Proc. Natl. Acad. Sci. USA 89:5907-5911, 1992; Primitive hematopoietic colony forming cells with high proliferative potential, McNiece, I.K. and Briddell, R.A. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 23-39, Wiley-Liss, Inc., New York, NY. 1994; Neben et al., Experimental Hematology 22:353-359, 1994; Cobblestone area forming cell assay, Ploemacher, R.E. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 1-21, Wiley-Liss, Inc.., New York, NY. 1994; Long term bone marrow cultures in the presence of stromal cells, Spooncer, E., Dexter, M. and Allen, T. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 163-179, Wiley-Liss, Inc., New York, NY. 1994; Long term culture initiating cell assay, Sutherland,

H.J. In *Culture of Hematopoietic Cells*. R.I. Freshney, *et al.* eds. Vol pp. 139-162, Wiley-Liss, Inc., New York, NY. 1994.

Tissue Growth Activity

A protein of the present invention also may have utility in compositions used for bone, cartilage, tendon, ligament and/or nerve tissue growth or regeneration, as well as for wound healing and tissue repair and replacement, and in the treatment of burns, incisions and ulcers.

A protein of the present invention, which induces cartilage and/or bone growth in circumstances where bone is not normally formed, has application in the healing of bone fractures and cartilage damage or defects in humans and other animals. Such a preparation employing a protein of the invention may have prophylactic use in closed as well as open fracture reduction and also in the improved fixation of artificial joints. *De novo* bone formation induced by an osteogenic agent contributes to the repair of congenital, trauma induced, or oncologic resection induced craniofacial defects, and also is useful in cosmetic plastic surgery.

A protein of this invention may also be used in the treatment of periodontal disease, and in other tooth repair processes. Such agents may provide an environment to attract bone-forming cells, stimulate growth of bone-forming cells or induce differentiation of progenitors of bone-forming cells. A protein of the invention may also be useful in the treatment of osteoporosis or osteoarthritis, such as through stimulation of bone and/or cartilage repair or by blocking inflammation or processes of tissue destruction (collagenase activity, osteoclast activity, etc.) mediated by inflammatory processes.

Another category of tissue regeneration activity that may be attributable to the protein of the present invention is tendon/ligament formation. A protein of the present invention, which induces tendon/ligament-like tissue or other tissue formation in circumstances where such tissue is not normally formed, has application in the healing of tendon or ligament tears, deformities and other tendon or ligament defects in humans and other animals. Such a preparation employing a tendon/ligament-like tissue inducing protein may have prophylactic use in preventing damage to tendon or ligament tissue, as well as use in the improved fixation of tendon or ligament to bone or other tissues, and in repairing defects to tendon or ligament tissue. De novo tendon/ligament-like tissue formation induced by a composition of the present invention contributes to the repair of

congenital, trauma induced, or other tendon or ligament defects of other origin, and is also useful in cosmetic plastic surgery for attachment or repair of tendons or ligaments. The compositions of the present invention may provide an environment to attract tendon- or ligament-forming cells, stimulate growth of tendon- or ligament-forming cells, induce differentiation of progenitors of tendon- or ligament-forming cells, or induce growth of tendon/ligament cells or progenitors *ex vivo* for return *in vivo* to effect tissue repair. The compositions of the invention may also be useful in the treatment of tendinitis, carpal tunnel syndrome and other tendon or ligament defects. The compositions may also include an appropriate matrix and/or sequestering agent as a carrier as is well known in the art.

The protein of the present invention may also be useful for proliferation of neural cells and for regeneration of nerve and brain tissue, *i.e.* for the treatment of central and peripheral nervous system diseases and neuropathies, as well as mechanical and traumatic disorders, which involve degeneration, death or trauma to neural cells or nerve tissue. More specifically, a protein may be used in the treatment of diseases of the peripheral nervous system, such as peripheral nerve injuries, peripheral neuropathy and localized neuropathies, and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager syndrome. Further conditions which may be treated in accordance with the present invention include mechanical and traumatic disorders, such as spinal cord disorders, head trauma and cerebrovascular diseases such as stroke. Peripheral neuropathies resulting from chemotherapy or other medical therapies may also be treatable using a protein of the invention.

Proteins of the invention may also be useful to promote better or faster closure of non-healing wounds, including without limitation pressure ulcers, ulcers associated with vascular insufficiency, surgical and traumatic wounds, and the like.

It is expected that a protein of the present invention may also exhibit activity for generation or regeneration of other tissues, such as organs (including, for example, pancreas, liver, intestine, kidney, skin, endothelium), muscle (smooth, skeletal or cardiac) and vascular (including vascular endothelium) tissue, or for promoting the growth of cells comprising such tissues. Part of the desired effects may be by inhibition or modulation of fibrotic scarring to allow normal tissue to regenerate. A protein of the invention may also exhibit angiogenic activity.

A protein of the present invention may also be useful for gut protection or regeneration and treatment of lung or liver fibrosis, reperfusion injury in various tissues, and conditions resulting from systemic cytokine damage.

A protein of the present invention may also be useful for promoting or inhibiting differentiation of tissues described above from precursor tissues or cells; or for inhibiting the growth of tissues described above.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for tissue generation activity include, without limitation, those described in: International Patent Publication No. WO95/16035 (bone, cartilage, tendon); International Patent Publication No. WO95/05846 (nerve, neuronal); International Patent Publication No. WO91/07491 (skin, endothelium).

Assays for wound healing activity include, without limitation, those described in: Winter, Epidermal Wound Healing, pps. 71-112 (Maibach, HI and Rovee, DT, eds.), Year Book Medical Publishers, Inc., Chicago, as modified by Eaglstein and Mertz, J. Invest. Dermatol 71:382-84 (1978).

Activin/Inhibin Activity

A protein of the present invention may also exhibit activin- or inhibin-related activities. Inhibins are characterized by their ability to inhibit the release of follicle stimulating hormone (FSH), while activins and are characterized by their ability to stimulate the release of follicle stimulating hormone (FSH). Thus, a protein of the present invention, alone or in heterodimers with a member of the inhibin α family, may be useful as a contraceptive based on the ability of inhibins to decrease fertility in female mammals and decrease spermatogenesis in male mammals. Administration of sufficient amounts of other inhibins can induce infertility in these mammals. Alternatively, the protein of the invention, as a homodimer or as a heterodimer with other protein subunits of the inhibin-β group, may be useful as a fertility inducing therapeutic, based upon the ability of activin molecules in stimulating FSH release from cells of the anterior pituitary. See, for example, United States Patent 4,798,885. A protein of the invention may also be useful for advancement of the onset of fertility in sexually immature mammals, so as to increase the lifetime reproductive performance of domestic animals such as cows, sheep and pigs.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for activin/inhibin activity include, without limitation, those described in: Vale et al., Endocrinology 91:562-572, 1972; Ling et al., Nature 321:779-782, 1986; Vale et al., Nature 321:776-779, 1986; Mason et al., Nature 318:659-663, 1985; Forage et al., Proc. Natl. Acad. Sci. USA 83:3091-3095, 1986.

Chemotactic/Chemokinetic Activity

A protein of the present invention may have chemotactic or chemokinetic activity (e.g., act as a chemokine) for mammalian cells, including, for example, monocytes, fibroblasts, neutrophils, T-cells, mast cells, eosinophils, epithelial and/or endothelial cells. Chemotactic and chemokinetic proteins can be used to mobilize or attract a desired cell population to a desired site of action. Chemotactic or chemokinetic proteins provide particular advantages in treatment of wounds and other trauma to tissues, as well as in treatment of localized infections. For example, attraction of lymphocytes, monocytes or neutrophils to tumors or sites of infection may result in improved immune responses against the tumor or infecting agent.

A protein or peptide has chemotactic activity for a particular cell population if it can stimulate, directly or indirectly, the directed orientation or movement of such cell population. Preferably, the protein or peptide has the ability to directly stimulate directed movement of cells. Whether a particular protein has chemotactic activity for a population of cells can be readily determined by employing such protein or peptide in any known assay for cell chemotaxis.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assays for chemotactic activity (which will identify proteins that induce or prevent chemotaxis) consist of assays that measure the ability of a protein to induce the migration of cells across a membrane as well as the ability of a protein to induce the adhesion of one cell population to another cell population. Suitable assays for movement and adhesion include, without limitation, those described in: Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and Wiley-Interscience (Chapter 6.12, Measurement of alpha and beta Chemokines 6.12.1-6.12.28; Taub et al. J. Clin. Invest. 95:1370-1376, 1995; Lind et al. APMIS 103:140-146, 1995; Muller et al Eur. J. Immunol. 25: 1744-1748; Gruber et al. J. of Immunol. 152:5860-5867, 1994; Johnston et al. J. of Immunol. 153: 1762-1768, 1994.

Hemostatic and Thrombolytic Activity

A protein of the invention may also exhibit hemostatic or thrombolytic activity. As a result, such a protein is expected to be useful in treatment of various coagulation disorders (including hereditary disorders, such as hemophilias) or to enhance coagulation and other hemostatic events in treating wounds resulting from trauma, surgery or other causes. A protein of the invention may also be useful for dissolving or inhibiting formation of thromboses and for treatment and prevention of conditions resulting therefrom (such as, for example, infarction of cardiac and central nervous system vessels (e.g., stroke).

The activity of a protein of the invention may, among other means, be measured by the following methods:

Assay for hemostatic and thrombolytic activity include, without limitation, those described in: Linet et al., J. Clin. Pharmacol. 26:131-140, 1986; Burdick et al., Thrombosis Res. 45:413-419, 1987; Humphrey et al., Fibrinolysis 5:71-79 (1991); Schaub, Prostaglandins 35:467-474, 1988.

Receptor/Ligand Activity

A protein of the present invention may also demonstrate activity as receptors, receptor ligands or inhibitors or agonists of receptor/ligand interactions. Examples of such receptors and ligands include, without limitation, cytokine receptors and their ligands, receptor kinases and their ligands, receptor phosphatases and their ligands, receptors involved in cell-cell interactions and their ligands (including without limitation, cellular adhesion molecules (such as selectins, integrins and their ligands) and receptor/ligand pairs involved in antigen presentation, antigen recognition and development of cellular and humoral immune responses). Receptors and ligands are also useful for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction. A protein of the present invention (including, without limitation, fragments of receptors and ligands) may themselves be useful as inhibitors of receptor/ligand interactions.

The activity of a protein of the invention may, among other means, be measured by the following methods:

Suitable assays for receptor-ligand activity include without limitation those described in:Current Protocols in Immunology, Ed by J.E. Coligan, A.M. Kruisbeek, D.H. Margulies, E.M. Shevach, W.Strober, Pub. Greene Publishing Associates and

Wiley-Interscience (Chapter 7.28, Measurement of Cellular Adhesion under static conditions 7.28.1-7.28.22), Takai et al., Proc. Natl. Acad. Sci. USA 84:6864-6868, 1987; Bierer et al., J. Exp. Med. 168:1145-1156, 1988; Rosenstein et al., J. Exp. Med. 169:149-160 1989; Stoltenborg et al., J. Immunol. Methods 175:59-68, 1994; Stitt et al., Cell 80:661-670, 1995.

Anti-Inflammatory Activity

Proteins of the present invention may also exhibit anti-inflammatory activity. The anti-inflammatory activity may be achieved by providing a stimulus to cells involved in the inflammatory response, by inhibiting or promoting cell-cell interactions (such as, for example, cell adhesion), by inhibiting or promoting chemotaxis of cells involved in the inflammatory process, inhibiting or promoting cell extravasation, or by stimulating or suppressing production of other factors which more directly inhibit or promote an inflammatory response. Proteins exhibiting such activities can be used to treat inflammatory conditions including chronic or acute conditions), including without limitation inflammation associated with infection (such as septic shock, sepsis or systemic inflammatory response syndrome (SIRS)), ischemia-reperfusion injury, endotoxin lethality, arthritis, complement-mediated hyperacute rejection, nephritis, cytokine or chemokine-induced lung injury, inflammatory bowel disease, Crohn's disease or resulting from over production of cytokines such as TNF or IL-1. Proteins of the invention may also be useful to treat anaphylaxis and hypersensitivity to an antigenic substance or material.

Cadherin/Tumor Invasion Suppressor Activity

Cadherins are calcium-dependent adhesion molecules that appear to play major roles during development, particularly in defining specific cell types. Loss or alteration of normal cadherin expression can lead to changes in cell adhesion properties linked to tumor growth and metastasis. Cadherin malfunction is also implicated in other human diseases, such as pemphigus vulgaris and pemphigus foliaceus (auto-immune blistering skin diseases), Crohn's disease, and some developmental abnormalities.

The cadherin superfamily includes well over forty members, each with a distinct pattern of expression. All members of the superfamily have in common conserved extracellular repeats (cadherin domains), but structural differences are found in other parts of the molecule. The cadherin domains bind calcium to form their tertiary structure and thus calcium is required to mediate their adhesion. Only a few amino acids in the

first cadherin domain provide the basis for homophilic adhesion; modification of this recognition site can change the specificity of a cadherin so that instead of recognizing only itself, the mutant molecule can now also bind to a different cadherin. In addition, some cadherins engage in heterophilic adhesion with other cadherins.

E-cadherin, one member of the cadherin superfamily, is expressed in epithelial cell types. Pathologically, if E-cadherin expression is lost in a tumor, the malignant cells become invasive and the cancer metastasizes. Transfection of cancer cell lines with polynucleotides expressing E-cadherin has reversed cancer-associated changes by returning altered cell shapes to normal, restoring cells' adhesiveness to each other and to their substrate, decreasing the cell growth rate, and drastically reducing anchorage-independent cell growth. Thus, reintroducing E-cadherin expression reverts carcinomas to a less advanced stage. It is likely that other cadherins have the same invasion suppressor role in carcinomas derived from other tissue types. Therefore, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be used to treat cancer. Introducing such proteins or polynucleotides into cancer cells can reduce or eliminate the cancerous changes observed in these cells by providing normal cadherin expression.

Cancer cells have also been shown to express cadherins of a different tissue type than their origin, thus allowing these cells to invade and metastasize in a different tissue in the body. Proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can be substituted in these cells for the inappropriately expressed cadherins, restoring normal cell adhesive properties and reducing or eliminating the tendency of the cells to metastasize.

Additionally, proteins of the present invention with cadherin activity, and polynucleotides of the present invention encoding such proteins, can used to generate antibodies recognizing and binding to cadherins. Such antibodies can be used to block the adhesion of inappropriately expressed tumor-cell cadherins, preventing the cells from forming a tumor elsewhere. Such an anti-cadherin antibody can also be used as a marker for the grade, pathological type, and prognosis of a cancer, i.e. the more progressed the cancer, the less cadherin expression there will be, and this decrease in cadherin expression can be detected by the use of a cadherin-binding antibody.

Fragments of proteins of the present invention with cadherin activity, preferably a polypeptide comprising a decapeptide of the cadherin recognition site, and polynucleotides of the present invention encoding such protein fragments, can also be used

to block cadherin function by binding to cadherins and preventing them from binding in ways that produce undesirable effects. Additionally, fragments of proteins of the present invention with cadherin activity, preferably truncated soluble cadherin fragments which have been found to be stable in the circulation of cancer patients, and polynucleotides encoding such protein fragments, can be used to disturb proper cell-cell adhesion.

Assays for cadherin adhesive and invasive suppressor activity include, without limitation, those described in: Hortsch et al. J Biol Chem 270 (32): 18809-18817, 1995; Miyaki et al. Oncogene 11: 2547-2552, 1995; Ozawa et al. Cell 63: 1033-1038, 1990.

Tumor Inhibition Activity

In addition to the activities described above for immunological treatment or prevention of tumors, a protein of the invention may exhibit other anti-tumor activities. A protein may inhibit tumor growth directly or indirectly (such as, for example, via ADCC). A protein may exhibit its tumor inhibitory activity by acting on tumor tissue or tumor precursor tissue, by inhibiting formation of tissues necessary to support tumor growth (such as, for example, by inhibiting angiogenesis), by causing production of other factors, agents or cell types which inhibit tumor growth, or by suppressing, eliminating or inhibiting factors, agents or cell types which promote tumor growth.

Other Activities

A protein of the invention may also exhibit one or more of the following additional activities or effects: inhibiting the growth, infection or function of, or killing, infectious agents, including, without limitation, bacteria, viruses, fungi and other parasites; effecting (suppressing or enhancing) bodily characteristics, including, without limitation, height, weight, hair color, eye color, skin, fat to lean ratio or other tissue pigmentation, or organ or body part size or shape (such as, for example, breast augmentation or diminution, change in bone form or shape); effecting biorhythms or caricadic cycles or rhythms; effecting the fertility of male or female subjects; effecting the metabolism, catabolism, anabolism, processing, utilization, storage or elimination of dietary fat, lipid, protein, carbohydrate, vitamins, minerals, cofactors or other nutritional factors or component(s); effecting behavioral characteristics, including, without limitation, appetite, libido, stress, cognition (including cognitive disorders), depression (including depressive disorders) and violent behaviors; providing analgesic effects or other pain reducing effects; promoting differentiation and growth of embryonic stem cells in lineages other than hematopoietic

lineages; hormonal or endocrine activity; in the case of enzymes, correcting deficiencies of the enzyme and treating deficiency-related diseases; treatment of hyperproliferative disorders (such as, for example, psoriasis); immunoglobulin-like activity (such as, for example, the ability to bind antigens or complement); and the ability to act as an antigen in a vaccine composition to raise an immune response against such protein or another material or entity which is cross-reactive with such protein.

ADMINISTRATION AND DOSING

A protein of the present invention (from whatever source derived, including without limitation from recombinant and non-recombinant sources) may be used in a pharmaceutical composition when combined with a pharmaceutically acceptable carrier. Such a composition may also contain (in addition to protein and a carrier) diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials well known in the art. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredient(s). The characteristics of the carrier will depend on the route of administration. The pharmaceutical composition of the invention may also contain cytokines, lymphokines, or other hematopoietic factors such as M-CSF, GM-CSF, TNF, IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-11, IL-12, IL-13, IL-14, IL-15, IFN, TNF0, TNF1, TNF2, G-CSF, Meg-CSF, thrombopoietin, stem cell factor, and erythropoietin. The pharmaceutical composition may further contain other agents which either enhance the activity of the protein or compliment its activity or use in treatment. Such additional factors and/or agents may be included in the pharmaceutical composition to produce a synergistic effect with protein of the invention, or to minimize side effects. Conversely, protein of the present invention may be included in formulations of the particular cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent to minimize side effects of the cytokine, lymphokine, other hematopoietic factor, thrombolytic or anti-thrombotic factor, or anti-inflammatory agent.

A protein of the present invention may be active in multimers (e.g., heterodimers or homodimers) or complexes with itself or other proteins. As a result, pharmaceutical compositions of the invention may comprise a protein of the invention in such multimeric or complexed form.

The pharmaceutical composition of the invention may be in the form of a complex of the protein(s) of present invention along with protein or peptide antigens. The protein and/or peptide antigen will deliver a stimulatory signal to both B and T lymphocytes. B lymphocytes will respond to antigen through their surface immunoglobulin receptor. T lymphocytes will respond to antigen through the T cell receptor (TCR) following presentation of the antigen by MHC proteins. MHC and structurally related proteins including those encoded by class I and class II MHC genes on host cells will serve to present the peptide antigen(s) to T lymphocytes. The antigen components could also be supplied as purified MHC-peptide complexes alone or with co-stimulatory molecules that can directly signal T cells. Alternatively antibodies able to bind surface immunolgobulin and other molecules on B cells as well as antibodies able to bind the TCR and other molecules on T cells can be combined with the pharmaceutical composition of the invention.

The pharmaceutical composition of the invention may be in the form of a liposome in which protein of the present invention is combined, in addition to other pharmaceutically acceptable carriers, with amphipathic agents such as lipids which exist in aggregated form as micelles, insoluble monolayers, liquid crystals, or lamellar layers in aqueous solution. Suitable lipids for liposomal formulation include, without limitation, monoglycerides, diglycerides, sulfatides, lysolecithin, phospholipids, saponin, bile acids, and the like. Preparation of such liposomal formulations is within the level of skill in the art, as disclosed, for example, in U.S. Patent No. 4,235,871; U.S. Patent No. 4,501,728; U.S. Patent No. 4,837,028; and U.S. Patent No. 4,737,323, all of which are incorporated herein by reference.

As used herein, the term "therapeutically effective amount" means the total amount of each active component of the pharmaceutical composition or method that is sufficient to show a meaningful patient benefit, i.e., treatment, healing, prevention or amelioration of the relevant medical condition, or an increase in rate of treatment, healing, prevention or amelioration of such conditions. When applied to an individual active ingredient, administered alone, the term refers to that ingredient alone. When applied to a combination, the term refers to combined amounts of the active ingredients that result in the therapeutic effect, whether administered in combination, serially or simultaneously.

In practicing the method of treatment or use of the present invention, a therapeutically effective amount of protein of the present invention is administered to a mammal having a condition to be treated. Protein of the present invention may be

administered in accordance with the method of the invention either alone or in combination with other therapies such as treatments employing cytokines, lymphokines or other hematopoietic factors. When co-administered with one or more cytokines, lymphokines or other hematopoietic factors, protein of the present invention may be administered either simultaneously with the cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors, or sequentially. If administered sequentially, the attending physician will decide on the appropriate sequence of administering protein of the present invention in combination with cytokine(s), lymphokine(s), other hematopoietic factor(s), thrombolytic or anti-thrombotic factors.

Administration of protein of the present invention used in the pharmaceutical composition or to practice the method of the present invention can be carried out in a variety of conventional ways, such as oral ingestion, inhalation, topical application or cutaneous, subcutaneous, intraperitoneal, parenteral or intravenous injection. Intravenous administration to the patient is preferred.

When a therapeutically effective amount of protein of the present invention is administered orally, protein of the present invention will be in the form of a tablet, capsule, powder, solution or elixir. When administered in tablet form, the pharmaceutical composition of the invention may additionally contain a solid carrier such as a gelatin or an adjuvant. The tablet, capsule, and powder contain from about 5 to 95% protein of the present invention, and preferably from about 25 to 90% protein of the present invention. When administered in liquid form, a liquid carrier such as water, petroleum, oils of animal or plant origin such as peanut oil, mineral oil, soybean oil, or sesame oil, or synthetic oils may be added. The liquid form of the pharmaceutical composition may further contain physiological saline solution, dextrose or other saccharide solution, or glycols such as ethylene glycol, propylene glycol or polyethylene glycol. When administered in liquid form, the pharmaceutical composition contains from about 0.5 to 90% by weight of protein of the present invention, and preferably from about 1 to 50% protein of the present invention.

When a therapeutically effective amount of protein of the present invention is administered by intravenous, cutaneous or subcutaneous injection, protein of the present invention will be in the form of a pyrogen-free, parenterally acceptable aqueous solution. The preparation of such parenterally acceptable protein solutions, having due regard to pH, isotonicity, stability, and the like, is within the skill in the art. A preferred

pharmaceutical composition for intravenous, cutaneous, or subcutaneous injection should contain, in addition to protein of the present invention, an isotonic vehicle such as Sodium Chloride Injection, Ringer's Injection, Dextrose Injection, Dextrose and Sodium Chloride Injection, Lactated Ringer's Injection, or other vehicle as known in the art. The pharmaceutical composition of the present invention may also contain stabilizers, preservatives, buffers, antioxidants, or other additives known to those of skill in the art.

The amount of protein of the present invention in the pharmaceutical composition of the present invention will depend upon the nature and severity of the condition being treated, and on the nature of prior treatments which the patient has undergone. Ultimately, the attending physician will decide the amount of protein of the present invention with which to treat each individual patient. Initially, the attending physician will administer low doses of protein of the present invention and observe the patient's response. Larger doses of protein of the present invention may be administered until the optimal therapeutic effect is obtained for the patient, and at that point the dosage is not increased further. It is contemplated that the various pharmaceutical compositions used to practice the method of the present invention should contain about 0.01 µg to about 100 mg (preferably about 0.1ng to about 10 mg, more preferably about 0.1 µg to about 1 mg) of protein of the present invention per kg body weight.

The duration of intravenous therapy using the pharmaceutical composition of the present invention will vary, depending on the severity of the disease being treated and the condition and potential idiosyncratic response of each individual patient. It is contemplated that the duration of each application of the protein of the present invention will be in the range of 12 to 24 hours of continuous intravenous administration. Ultimately the attending physician will decide on the appropriate duration of intravenous therapy using the pharmaceutical composition of the present invention.

Protein of the invention may also be used to immunize animals to obtain polyclonal and monoclonal antibodies which specifically react with the protein. Such antibodies may be obtained using either the entire protein or fragments thereof as an immunogen. The peptide immunogens additionally may contain a cysteine residue at the carboxyl terminus, and are conjugated to a hapten such as keyhole limpet hemocyanin (KLH). Methods for synthesizing such peptides are known in the art, for example, as in R.P. Merrifield, J. Amer.Chem.Soc. 85, 2149-2154 (1963); J.L. Krstenansky, *et al.*, FEBS Lett. 211, 10 (1987). Monoclonal antibodies binding to the protein of the invention may be useful diagnostic agents for the immunodetection of the protein. Neutralizing monoclonal

antibodies binding to the protein may also be useful therapeutics for both conditions associated with the protein and also in the treatment of some forms of cancer where abnormal expression of the protein is involved. In the case of cancerous cells or leukemic cells, neutralizing monoclonal antibodies against the protein may be useful in detecting and preventing the metastatic spread of the cancerous cells, which may be mediated by the protein.

For compositions of the present invention which are useful for bone, cartilage, tendon or ligament regeneration, the therapeutic method includes administering the composition topically, systematically, or locally as an implant or device. When administered, the therapeutic composition for use in this invention is, of course, in a pyrogen-free, physiologically acceptable form. Further, the composition may desirably be encapsulated or injected in a viscous form for delivery to the site of bone, cartilage or tissue damage. Topical administration may be suitable for wound healing and tissue repair. Therapeutically useful agents other than a protein of the invention which may also optionally be included in the composition as described above, may alternatively or additionally, be administered simultaneously or sequentially with the composition in the methods of the invention. Preferably for bone and/or cartilage formation, the composition would include a matrix capable of delivering the protein-containing composition to the site of bone and/or cartilage damage, providing a structure for the developing bone and cartilage and optimally capable of being resorbed into the body. Such matrices may be formed of materials presently in use for other implanted medical applications.

The choice of matrix material is based on biocompatibility, biodegradability, mechanical properties, cosmetic appearance and interface properties. The particular application of the compositions will define the appropriate formulation. Potential matrices for the compositions may be biodegradable and chemically defined calcium sulfate, tricalciumphosphate, hydroxyapatite, polylactic acid, polyglycolic acid and polyanhydrides. Other potential materials are biodegradable and biologically well-defined, such as bone or dermal collagen. Further matrices are comprised of pure proteins or extracellular matrix components. Other potential matrices are nonbiodegradable and chemically defined, such as sintered hydroxapatite, bioglass, aluminates, or other ceramics. Matrices may be comprised of combinations of any of the above mentioned types of material, such as polylactic acid and hydroxyapatite or collagen and tricalciumphosphate. The bioceramics may be altered in composition, such as in calcium-

aluminate-phosphate and processing to alter pore size, particle size, particle shape, and biodegradability.

Presently preferred is a 50:50 (mole weight) copolymer of lactic acid and glycolic acid in the form of porous particles having diameters ranging from 150 to 800 microns. In some applications, it will be useful to utilize a sequestering agent, such as carboxymethyl cellulose or autologous blood clot, to prevent the protein compositions from disassociating from the matrix.

A preferred family of sequestering agents is cellulosic materials such as alkylcelluloses (including hydroxyalkylcelluloses), including methylcellulose, ethylcellulose, hydroxyethylcellulose, hydroxypropylcellulose, hydroxypropyl-methylcellulose, and carboxymethylcellulose, the most preferred being cationic salts of carboxymethylcellulose (CMC). Other preferred sequestering agents include hyaluronic acid, sodium alginate, poly(ethylene glycol), polyoxyethylene oxide, carboxyvinyl polymer and poly(vinyl alcohol). The amount of sequestering agent useful herein is 0.5-20 wt%, preferably 1-10 wt% based on total formulation weight, which represents the amount necessary to prevent desorbtion of the protein from the polymer matrix and to provide appropriate handling of the composition, yet not so much that the progenitor cells are prevented from infiltrating the matrix, thereby providing the protein the opportunity to assist the osteogenic activity of the progenitor cells.

In further compositions, proteins of the invention may be combined with other agents beneficial to the treatment of the bone and/or cartilage defect, wound, or tissue in question. These agents include various growth factors such as epidermal growth factor (EGF), platelet derived growth factor (PDGF), transforming growth factors (TGF-α and TGF-β), and insulin-like growth factor (IGF).

The therapeutic compositions are also presently valuable for veterinary applications. Particularly domestic animals and thoroughbred horses, in addition to humans, are desired patients for such treatment with proteins of the present invention.

The dosage regimen of a protein-containing pharmaceutical composition to be used in tissue regeneration will be determined by the attending physician considering various factors which modify the action of the proteins, e.g., amount of tissue weight desired to be formed, the site of damage, the condition of the damaged tissue, the size of a wound, type of damaged tissue (e.g., bone), the patient's age, sex, and diet, the severity of any infection, time of administration and other clinical factors. The dosage may vary with the type of matrix used in the reconstitution and with inclusion of other proteins in

the pharmaceutical composition. For example, the addition of other known growth factors, such as IGF I (insulin like growth factor I), to the final composition, may also effect the dosage. Progress can be monitored by periodic assessment of tissue/bone growth and/or repair, for example, X-rays, histomorphometric determinations and tetracycline labeling.

Polynucleotides of the present invention can also be used for gene therapy. Such polynucleotides can be introduced either *in vivo* or *ex vivo* into cells for expression in a mammalian subject. Polynucleotides of the invention may also be administered by other known methods for introduction of nucleic acid into a cell or organism (including, without limitation, in the form of viral vectors or naked DNA).

Cells may also be cultured *ex vivo* in the presence of proteins of the present invention in order to proliferate or to produce a desired effect on or activity in such cells. Treated cells can then be introduced *in vivo* for therapeutic purposes.

Patent and literature references cited herein are incorporated by reference as if fully set forth.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Jacobs, Kenneth
McCoy, John M.
LaVallie, Edward R.
Racie, Lisa A.
Merberg, David
Treacy, Maurice
Spaulding, Vikki
Agostino, Michael J.

(ii) TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM

(iii) NUMBER OF SEQUENCES: 25

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genetics Institute, Inc.
(B) STREET: 87 CambridgePark Drive
(C) CITY: Cambridge
(D) STATE: MA
(E) COUNTRY: U.S.A.
(F) ZIP: 02140

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Sprunger, Suzanne A.
(B) REGISTRATION NUMBER: 41,323

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617) 498-8284
(B) TELEFAX: (617) 876-5851

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1480 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AGGCGCCCTC CCTTCCTGAG GAGCTGTTGG CCTGGGTGGG CAGGAACTGC AGTATGGCCA 60

TGGGCTGAGC AGGCTGAGCA CCTCAGCCTT TAGGGCTTAT GGCCAGGGGA CACTGTATGA 120

CTCTCCTCTC CTGCAGGTGT CTATCCACCT GGGGTATGGC ATCTACCGAC CTGTCTCCCT 180

GGGGTCACAT GCTTTGTTTC CATTCTTGTC CTGGCTGGAC CAGCCACTGT GGGACCAACA 240

CCCCTCCCAC ACTCCCCCAG ACTGCTCGTC TATCACCAGG ATCGCTTTGT ACTTTGTGCA 300

AAAGGGTCTG GCTGTCCCTT GCTGTTTTCA TCTCTGCCAA GCCTATTGTG CCTCTGGCTG 360

CTGTATGTGT GCGCGTGCAC GTGTGTGTGT TTCATCTGTT CATTCACTGC ACAAGATATT 420

TATTGAGTGC CCACTACGTG CCAGGCACTG TTGCTGAGTT CCTGTGGGTG TGTCTCTCGA 480

TGCCACTCCT GCTTCTCTGG GGGCCTCTTT CTGTGCTTCT CTTTGTCCCC AAATTGCTAC 540

CTCTTTGTCA GTCTGGGTGT CTCAGGTTCT GTGTGTCCTT GTGTGCATTT CTGTCTCTCT 600

CTGTCCTCGT CTCTCTGCAA GGCCCTCTAT TTCTCTCTTT CTTGGTGTCT GTCCTTTGCC 660

CCCTGTGCCC TCTGGATTCT CTGGGTCTAT GTAGGCCCCT GGTCTGCCCT GGGCTCATCA 720

GCCTTCCTGA CCTCCTCCTG CCCTCCCCTT CACTCCCTCC CTGGCTCTGC CAGTCGGTTC 780

CCACGGAGCC ATTTTTAGCT CTGATCAGCA TGGGAATGTG CCTCGGCCTC CAAGGGGCTT 840

TGTCCTGGTG CCCCCGCCCC TGGTCCCAAC CTGATCCCAC GAGGGAGTTG GGACAGGAGG 900

ATTGATGGTG CTCCCCTTCC TGCCAGCGTC AGAGGCCCTG GAGAGGGGCT GTCCATGGCA 960

GCTGGTCTTT ATTCCTCCCT CATGAGCACA GGGTCGGGGG GTCCCCATTC TTGGAAGAGG 1020

TTGAGAAGAC TCCTGGGCTT CAGCCTCTCC CACCCAGCCC TGCCCCTCAC CTGCCTGCCC 1080

TCCCCTCCCC CACTCTATAC TAGGGACTGG ATCTCAGCCT CTGATCAGTT TCACAAAGTT 1140

TGTTCCCTAA GGAAATCAAA TCCCATTGTC ACCTAACTCT GAAGATCTAA ATAGCCCTTG 1200

GATCAGTACG GGAACCCCAA ATCCCACAGG GCCAGATGTG GAGTCTGTGT CTGCCCCCGT 1260

CTTCTCTCCA TCCTCAAAGC CCCCACTTCT CTCCAGGCTG TTTCTTTTTT TATGACTGTA 1320

AACATAGATA GTGCTTTATT TTGTTAATAA TAAGATAATG ATGAGTAACT TAACCAGCAC 1380

ATTTCTCCTG TTTACACTCG GGGGATTTTT TTGTTTTCTG ATGACATAAT AAAGACAGAT 1440

CATTTCAGAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1480

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 268 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Arg Gly His Cys Met Thr Leu Leu Ser Cys Arg Cys Leu Ser
1 5 10 15

Thr Trp Gly Met Ala Ser Thr Asp Leu Ser Pro Trp Gly His Met Leu
20 25 30

Cys Phe His Ser Cys Pro Gly Trp Thr Ser His Cys Gly Thr Asn Thr
35 40 45

Pro Pro Thr Leu Pro Gln Thr Ala Arg Leu Ser Pro Gly Ser Leu Cys
50 55 60

Thr Leu Cys Lys Arg Val Trp Leu Ser Leu Ala Val Phe Ile Ser Ala
65 70 75 80

Lys Pro Ile Val Pro Leu Ala Ala Val Cys Val Arg Val His Val Cys
85 90 95

Val Phe His Leu Phe Ile His Cys Thr Arg Tyr Leu Leu Ser Ala His
100 105 110

Tyr Val Pro Gly Thr Val Ala Glu Phe Leu Trp Val Cys Leu Ser Met
115 120 125

Pro Leu Leu Leu Leu Trp Gly Pro Leu Ser Val Leu Leu Phe Val Pro
130 135 140

Lys Leu Leu Pro Leu Cys Gln Ser Gly Cys Leu Arg Phe Cys Val Ser
145 150 155 160

Leu Cys Ala Phe Leu Ser Leu Ser Val Leu Val Ser Leu Gln Gly Pro
165 170 175

Leu Phe Leu Ser Phe Leu Val Ser Val Leu Cys Pro Leu Cys Pro Leu
180 185 190

Asp Ser Leu Gly Leu Cys Arg Pro Leu Val Cys Pro Gly Leu Ile Ser
195 200 205

Leu Pro Asp Leu Leu Leu Pro Ser Pro Ser Leu Pro Pro Trp Leu Cys
210 215 220

Gln Ser Val Pro Thr Glu Pro Phe Leu Ala Leu Ile Ser Met Gly Met
225 230 235 240

Cys Leu Gly Leu Gln Gly Ala Leu Ser Trp Cys Pro Arg Pro Trp Ser
245 250 255

Gln Pro Asp Pro Thr Arg Glu Leu Gly Gln Glu Asp
260 265

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1436 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CCCGGCGGCT CCTGGAACCC CGGTTCGCGG CGATGCCAGC CACCCCAGCG AAGCCGCCGC 60

AGTTCAGTGC TTGGATAATT TGAAAGTACA ATAGTTGGTT TCCCTGTCCA CCCGCCCCAC 120

TTCGCTTGCC ATCACAGCAC GCCTATCGGA TGTGAGAGGA GAAGTCCCGC TGCTCGGGCA 180

CTGTCTATAT ACGCCTAACA CCTACATATA TTTTAAAAAC ATTAAATATA ATTAACAATC 240

AAAAGAAAGA GGAGAAAGGA AGGGAAGCAT TACTGGGTTA CTATGCACTT GCGACTGATT 300

TCTTGGCTTT TTATCATTTT GAACTTTATG GAATACATCG GCAGCCAAAA CGCCTCCCGG 360

GGAAGGCGCC AGCGAAGAAT GCATCCTAAC GTTAGTCAAG GCTGCCAAGG AGGCTGTGCA 420

ACATGCTCAG ATTACAATGG ATGTTTGTCA TGTAAGCCCA GACTATTTTT TGCTCTGGAA 480

AGAATTGGCA TGAAGCAGAT TGGAGTATGT CTCTCTTCAT GTCCAAGTGG ATATTATGGA 540

ACTCGATATC CAGATATAAA TAAGTGTACA AAATGCAAAG CTGACTGTGA TACCTGTTTC 600

AACAAAAATT TCTGCACAAA ATGTAAAAGT GGATTTTACT TACACCTTGG AAAGTGCCTT 660

GACAATTGCC CAGAAGGGTT GGAAGCCAAC AACCATACTA TGGAGTGTGT CAGTATTGTG 720

CACTGTGAGG TCAGTGAATG GAATCCTTGG AGTCCATGCA CGAAGAAGGG AAAAACATGT 780

GGCTTCAAAA GAGGGACTGA AACACGGGTC CGAGAAATAA TACAGCATCC TTCAGCAAAG 840

GGTAACCTGT GTCCCCCAAC AAATGAGACA AGAAAGTGTA CAGTGCAAAG GAAGAAGTGT 900

CAGAAGGGAG AACGAGGAAA AAAAGGAAGG GAGAGGAAAA GAAAAAAACC TAATAAAGGA 960

GAAAGTAAAG AAGCAATACC TGACAGCAAA AGTCTGGAAT CCAGCAAAGA AATCCCAGAG 1020

CAACGAGAAA ACAAACAGCA GCAGAAGAAG CGAAAAGTCC AAGATAAACA GAAATCGGGG 1080

ATTGAAGTCA CCCTAGCTGA AGGCCTCACC AGTGTTTCAC AGAGGACACA GCCCACCCCT 1140

TGCAGGAGGA GGTATCTCTG AGTGTGCAGC ACAGAATCGC ATGACCCACC TTAACCTTCC 1200

TGTTGTCATG GAAGGATGCA CGGCTGCTCT GTCCACTGTG ATTCCTAGCC CTCTCAAGAT 1260

CACTGCTTTC TGAAGAATTT GCAATGACTC TGGCTTCTGG CTGCTTATCT CTGGACACCC 1320

GTTCTCCACC AGTTGTACAG TTCATGTAAT CTACTTGGCT TAATTGATTT TCCACTTCTC 1380

TCTTCCTCTT CTAAGATATA AACATTTTAA ATGATTTAAA AAAAAAAAAA AAAAAA 1436

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 292 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
Met His Leu Arg Leu Ile Ser Trp Leu Phe Ile Ile Leu Asn Phe Met
1               5                  10                  15

Glu Tyr Ile Gly Ser Gln Asn Ala Ser Arg Gly Arg Arg Gln Arg Arg
            20                  25                  30

Met His Pro Asn Val Ser Gln Gly Cys Gln Gly Gly Cys Ala Thr Cys
        35                  40                  45

Ser Asp Tyr Asn Gly Cys Leu Ser Cys Lys Pro Arg Leu Phe Phe Ala
    50                  55                  60

Leu Glu Arg Ile Gly Met Lys Gln Ile Gly Val Cys Leu Ser Ser Cys
65                  70                  75                  80

Pro Ser Gly Tyr Tyr Gly Thr Arg Tyr Pro Asp Ile Asn Lys Cys Thr
                85                  90                  95

Lys Cys Lys Ala Asp Cys Asp Thr Cys Phe Asn Lys Asn Phe Cys Thr
            100                 105                 110
```

```
Lys Cys Lys Ser Gly Phe Tyr Leu His Leu Gly Lys Cys Leu Asp Asn
        115                 120                 125

Cys Pro Glu Gly Leu Glu Ala Asn Asn His Thr Met Glu Cys Val Ser
    130                 135                 140

Ile Val His Cys Glu Val Ser Glu Trp Asn Pro Trp Ser Pro Cys Thr
145                 150                 155                 160

Lys Lys Gly Lys Thr Cys Gly Phe Lys Arg Gly Thr Glu Thr Arg Val
                165                 170                 175

Arg Glu Ile Ile Gln His Pro Ser Ala Lys Gly Asn Leu Cys Pro Pro
            180                 185                 190

Thr Asn Glu Thr Arg Lys Cys Thr Val Gln Arg Lys Lys Cys Gln Lys
        195                 200                 205

Gly Glu Arg Gly Lys Lys Gly Arg Glu Arg Lys Arg Lys Lys Pro Asn
    210                 215                 220

Lys Gly Glu Ser Lys Glu Ala Ile Pro Asp Ser Lys Ser Leu Glu Ser
225                 230                 235                 240

Ser Lys Glu Ile Pro Glu Gln Arg Glu Asn Lys Gln Gln Gln Lys Lys
                245                 250                 255

Arg Lys Val Gln Asp Lys Gln Lys Ser Gly Ile Glu Val Thr Leu Ala
            260                 265                 270

Glu Gly Leu Thr Ser Val Ser Gln Arg Thr Gln Pro Thr Pro Cys Arg
        275                 280                 285

Arg Arg Tyr Leu
    290
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2322 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
GGTTAAGAGC AGATTAGAAC AGAAATCAGG AGAACTTGGG AAGAAGATCA CTGAGTTAAC      60

ATTGAAAAAT CAGACACTAC AAAAGGAAAT TGAAAAAGTT TATTTGGATA ATAAGCTCCT     120
```

CAAGGAGCAA GCACATAACT TAACAATTGA AATGAAAAAT CATTATGTTC CTTTAAAAGT 180

AAGTGAAGAC ATGAAAAAGT CACATGATGC AATTATTGAT GATCTTAATA GAAAGCTTTT 240

AGATGTAACA CAAAAATATA CAGAAAAGAA GTTGGAAATG GAGAAATTGC TACTGGAAAA 300

TGACAGCTTA AGTAAGGATG TAAGCCGCCT AGAAACTGTG TTTGTACCTC CTGAGAAACA 360

TGAAAAAGAG ATAATAGCTC TGAAATCCAA TATTGTTGAA CTTAAGAAAC AGCTGTCTGA 420

ACTTAAGAAA AAATGTGGTG AAGACCAGGA GAAAATACAC GCTCTCACAT CTGAAAACAC 480

TAACTTGAAG AAGATGATGA GTAATCAGTA TGTGCCAGTT AAAACCCATG AAGAGGTTAA 540

AATGACACTG AATGACACGT TAGCCAAAAC TAACAGAGAA TTATTAGATG TGAAGAAAAA 600

ATTTGAAGAT ATAAATCAGG AATTTGTAAA AATAAAAGAT AAGAATGAAA TATTAAAAAG 660

AAACCTGGAA AACACTCAGA ACCAAATAAA AGCTGAGTAC ATCAGCCTGG CAGAGCACGA 720

GGCAAAGATG AGCTCGCTAA GTCAGAGCAT GAGAAAGGTG CAGGATAGTA ATGCTGAAAT 780

CTTGGCCAAC TACAGAAAAG GCCAAGAAGA GATTGTGACA CTGCATGCCG AAATTAAAGC 840

CCAGAAGAAG GAGCTCGACA CAATACAAGA ATGCATTAAG GTAAAATATG CCCCAATTGT 900

CAGCTTTGAG GAGTGCGAGA GAAAATTTAA AGCAACAGAG AAAGAACTAA AAGACCAGTT 960

ATCAGAGCAG ACACAAAAGT ATAGTGTCAG TGAAGAAGAA GTCAAGAAAA ACAAGCAAGA 1020

GAATGACAAG TTAAAGAAGG AGATTTTTAC CCTTCAGAAA GATTTGAGAG ATAAGACAGT 1080

TCTCATTGAG AAGTCTCATG AAATGGAAAG AGCATTAAGC AGAAAAACAG ACGAGCTAAA 1140

CAAACAGTTA AAAGACTTGT CACAGAAATA CACGGAAGTA AAGAATGTGA AAGAGAAGCT 1200

AGTAGAAGAA AATGCCAAAC AGACTTCTGA GATACTTGCA GTGCAAAATC TTTTGCAAAA 1260

ACAACATGTT CCATTGGAAC AGGTTGAGGC TCTGAAAAAA TCTCTTAATG GCACAATTGA 1320

AAATCTAAAG GAAGAACTGA AGAGTATGCA AAGGTGTTAC GAGAAAGAGC AGCAGACAGT 1380

GACCAAACTG CATCAATTGT TGGAGAATCA AAAGAACTCT TCTGTACCCC TGGCAGAGCA 1440

TTTGCAGATT AAAGAAGCAT TTGAGAAAGA AGTTGGAATC ATAAAAGCCA GCTTGAGAGA 1500

AAAGGAAGAA GAAAGCCAAA ACAAAATGGA AGAAGTCTCC AAACTTCAGT CGGAGGTTCA 1560

GAATACTAAA CAAGCATTAA AAAAATTAGA GACTAGAGAG GTAGTTGACT TGTCTAAATA 1620

TAAAGCAACA AAAAGTGATT TGGAGACACA GATTTCTAGC TTAAATGAAA AATTGGCCAA 1680

TCTGAATAGA AAGTATGAGG AAGTATGTGA GGAAGTTTTG CATGCCAAAA AGAAGGAAAT 1740

ATCTGCAAAA GATGAGAAGG AATTACTGCA TTTCAGCATT GAGCAAGAAA TTAAGGATCA 1800

GAAGGAACGA TGTGATAAGT CCTTAACAAC AATCACAGAG TTACAAAGAA GAATACAAGA 1860

ATCTGCTAAA CAAATAGAAG CAAAAGATAA TAAGATAACT GAACTGCTTA ATGATGTGGA 1920

AAGATTAAAA CAGGCACTCA ATGGCCTTTC CCAACTCACC TACACAAGTG GGAACCCCAC 1980

CAAGAGGCAG AGCCAGCTGA TTGACACTCT GCAGCACCAA GTGAAATCTC TGGAGCAACA 2040

GCTGGCCGAT GCTGACAGAC AGCACCAAGA AGTAATTGCA ATTTATCGGA CACACCTTCT 2100

TAGTGCTGCA CAGGGTCACA TGGATGAAGA TGTTCAGGAG GCTCTGCTCC AGATCATACA 2160

AATGCGGCAG GGGCTTGTGT GCTAGCCGTT AGCACTGACT GCCAGTATCT GTTTTATCTT 2220

GCTGGTGCTG AACATTCTTT GTGCAACTCC ATGGTCTTTC TGGGCCTTAC TGTGCTGGTA 2280

TAATTAAAAT AAAATATATT TTGTTCTAAA AAAAAAAAAA AA 2322

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 677 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Lys Asn His Tyr Val Pro Leu Lys Val Ser Glu Asp Met Lys Lys
1 5 10 15

Ser His Asp Ala Ile Ile Asp Asp Leu Asn Arg Lys Leu Leu Asp Val
20 25 30

Thr Gln Lys Tyr Thr Glu Lys Lys Leu Glu Met Glu Lys Leu Leu Leu
35 40 45

Glu Asn Asp Ser Leu Ser Lys Asp Val Ser Arg Leu Glu Thr Val Phe
50 55 60

Val Pro Pro Glu Lys His Glu Lys Glu Ile Ile Ala Leu Lys Ser Asn
65 70 75 80

Ile Val Glu Leu Lys Lys Gln Leu Ser Glu Leu Lys Lys Lys Cys Gly
85 90 95

Glu Asp Gln Glu Lys Ile His Ala Leu Thr Ser Glu Asn Thr Asn Leu
100 105 110

Lys Lys Met Met Ser Asn Gln Tyr Val Pro Val Lys Thr His Glu Glu

```
        115                 120                 125

Val Lys Met Thr Leu Asn Asp Thr Leu Ala Lys Thr Asn Arg Glu Leu
    130                 135                 140

Leu Asp Val Lys Lys Lys Phe Glu Asp Ile Asn Gln Glu Phe Val Lys
145                 150                 155                 160

Ile Lys Asp Lys Asn Glu Ile Leu Lys Arg Asn Leu Glu Asn Thr Gln
                165                 170                 175

Asn Gln Ile Lys Ala Glu Tyr Ile Ser Leu Ala Glu His Glu Ala Lys
            180                 185                 190

Met Ser Ser Leu Ser Gln Ser Met Arg Lys Val Gln Asp Ser Asn Ala
        195                 200                 205

Glu Ile Leu Ala Asn Tyr Arg Lys Gly Gln Glu Glu Ile Val Thr Leu
    210                 215                 220

His Ala Glu Ile Lys Ala Gln Lys Lys Glu Leu Asp Thr Ile Gln Glu
225                 230                 235                 240

Cys Ile Lys Val Lys Tyr Ala Pro Ile Val Ser Phe Glu Glu Cys Glu
                245                 250                 255

Arg Lys Phe Lys Ala Thr Glu Lys Glu Leu Lys Asp Gln Leu Ser Glu
            260                 265                 270

Gln Thr Gln Lys Tyr Ser Val Ser Glu Glu Glu Val Lys Lys Asn Lys
        275                 280                 285

Gln Glu Asn Asp Lys Leu Lys Lys Glu Ile Phe Thr Leu Gln Lys Asp
    290                 295                 300

Leu Arg Asp Lys Thr Val Leu Ile Glu Lys Ser His Glu Met Glu Arg
305                 310                 315                 320

Ala Leu Ser Arg Lys Thr Asp Glu Leu Asn Lys Gln Leu Lys Asp Leu
                325                 330                 335

Ser Gln Lys Tyr Thr Glu Val Lys Asn Val Lys Glu Lys Leu Val Glu
            340                 345                 350

Glu Asn Ala Lys Gln Thr Ser Glu Ile Leu Ala Val Gln Asn Leu Leu
        355                 360                 365

Gln Lys Gln His Val Pro Leu Glu Gln Val Glu Ala Leu Lys Lys Ser
    370                 375                 380

Leu Asn Gly Thr Ile Glu Asn Leu Lys Glu Glu Leu Lys Ser Met Gln
385                 390                 395                 400

Arg Cys Tyr Glu Lys Glu Gln Gln Thr Val Thr Lys Leu His Gln Leu
                405                 410                 415
```

```
Leu Glu Asn Gln Lys Asn Ser Ser Val Pro Leu Ala Glu His Leu Gln
            420                 425                 430

Ile Lys Glu Ala Phe Glu Lys Glu Val Gly Ile Ile Lys Ala Ser Leu
        435                 440                 445

Arg Glu Lys Glu Glu Glu Ser Gln Asn Lys Met Glu Glu Val Ser Lys
    450                 455                 460

Leu Gln Ser Glu Val Gln Asn Thr Lys Gln Ala Leu Lys Lys Leu Glu
465                 470                 475                 480

Thr Arg Glu Val Val Asp Leu Ser Lys Tyr Lys Ala Thr Lys Ser Asp
                485                 490                 495

Leu Glu Thr Gln Ile Ser Ser Leu Asn Glu Lys Leu Ala Asn Leu Asn
            500                 505                 510

Arg Lys Tyr Glu Glu Val Cys Glu Glu Val Leu His Ala Lys Lys Lys
        515                 520                 525

Glu Ile Ser Ala Lys Asp Glu Lys Glu Leu Leu His Phe Ser Ile Glu
    530                 535                 540

Gln Glu Ile Lys Asp Gln Lys Glu Arg Cys Asp Lys Ser Leu Thr Thr
545                 550                 555                 560

Ile Thr Glu Leu Gln Arg Arg Ile Gln Glu Ser Ala Lys Gln Ile Glu
                565                 570                 575

Ala Lys Asp Asn Lys Ile Thr Glu Leu Leu Asn Asp Val Glu Arg Leu
            580                 585                 590

Lys Gln Ala Leu Asn Gly Leu Ser Gln Leu Thr Tyr Thr Ser Gly Asn
        595                 600                 605

Pro Thr Lys Arg Gln Ser Gln Leu Ile Asp Thr Leu Gln His Gln Val
    610                 615                 620

Lys Ser Leu Glu Gln Gln Leu Ala Asp Ala Asp Arg Gln His Gln Glu
625                 630                 635                 640

Val Ile Ala Ile Tyr Arg Thr His Leu Leu Ser Ala Ala Gln Gly His
                645                 650                 655

Met Asp Glu Asp Val Gln Glu Ala Leu Leu Gln Ile Ile Gln Met Arg
            660                 665                 670

Gln Gly Leu Val Cys
        675
```

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2041 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

TCTCCCCCCT CCCCGACACA CACTCACAGG CCGGGCATTG ATGGTAATGT ATGCGAGGAA 60

ACAGCAGAGA CTCAGTGATG GCTGTCACGA CCGGAGGGGG GACTCGCAGC CTTACCAGGC 120

ACTTAAGTAT TCATCGAAGA GTCACCCCAG TAGCGGTGAT CACAGACATG AAAAGATGCG 180

AGACGCCGGA GATCCTTCAC CACCAAATAA AATGTTGCGG AGATCTGATA GTCCTGAAAA 240

CAAATACAGT GACAGCACAG GTCACAGTAA GGCCAAAAAT GTGCATACTC ACAGAGTTAG 300

AGAGAGGGAT GGTGGGACCA GTTACTCTCC ACAAGAAAAT TCACACAACC ACAGTGCTCT 360

TCATAGTTCA AATTCACATT CTTCTAATCC AAGCAATAAC CCAAGCAAAA CTTCAGATGC 420

ACCTTATGAT TCTGCAGATG ACTGGTCTGA GCATATTAGC TCTTCTGGGA AAAAGTACTA 480

CTACAATTGT CGAACAGAAG TTTCACAATG GGAAAAACCA AAAGAGTGGC TTGAAAGAGA 540

ACAGAGACAA AAAGAAGCAA ACAAGATGGC AGTCAACAGC TTCCCAAAAG ATAGGGATTA 600

CAGAAGAGAG GTGATGCAAG CAACAGCCAC TAGTGGGTTT GCCAGTGGAA AATCTACATC 660

AGGAGACAAA CCCGTATCAC ATTCTTGCAC AACTCCTTCC ACGTCTTCTG CCTCTGGACT 720

GAACCCCACA TCTGCACCTC CAACATCTGC TTCAGCGGTC CCTGTTTCTC CTGTTCCACA 780

GTCGCCAATA CCTCCCTTAC TTCAGGACCC AAATCTTCTT AGACAATTGC TTCCTGCTTT 840

GCAAGCCACG CTGCAGCTTA ATAATTCTAA TGTGGACATA TCTAAAATAA ATGAAGTTCT 900

TACAGCAGCT GTGACACAAG CCTCACTGCA GTCTATAATT CATAAGTTTC TTACTGCTGG 960

ACCATCTGCT TTCAACATAA CGTCTCTGAT TTCTCAAGCT GCTCAGCTCT CTACACAAGC 1020

CCAGCCATCT AATCAGTCTC CGATGTCTTT AACATCTGAT GCGTCATCCC CAAGATCATA 1080

TGTTTCTCCA AGAATAAGCA CACCTCAAAC TAACACAGTC CCTATCAAAC CTTTGATCAG 1140

TACTCCTCCT GTTTCATCAC AGCCAAAGGT TAGTACTCCA GTAGTTAAGC AAGGACCAGT 1200

GTCACAGTCA GCCACACAGC AGCCTGTAAC TGCTGACAAG CAGCAAGGTC ATGAACCTGT 1260

CTCTCCTCGA AGTCTTCAGC GCTCAAGCCA GAGAAGTCCA TCACCTGGTC CCAATCATAC 1320

TTCTAATAGT AGTAATGCAT CAAATGCAAC AGTTGTACCA CAGAATTCTT CTGCCCGATC 1380

CACGTGTTCA TTAACGCCTG CACTAGCAGC ACACTTCAGT GAAAATCTCA TAAAACACGT 1440

TCAAGGATGG CCTGCAGATC ATGCAGAGAA GCAGGCATCA AGATTACGCG AAGAAGCGCA 1500

TAACATGGGA ACTATTCACA TGTCCGAAAT TTGTACTGAA TTAAAAAATT TAAGATCTTT 1560

AGTCCGAGTA TGTGAAATTC AAGCAACTTT GCGAGAGCAA AGGATACTAT TTTTGAGACA 1620

ACAAATTAAG GAACTTGAAA AGCTAAAAAA TCAGAATTCC TTCATGGTGT GAAGATGTGA 1680

ATAATTGCAC ATGGTTTTGA GAACAGGAAC TGTAAATCTG TTGCCCAATC TTAACATTTT 1740

TGAGCTGCAT TTAAGTAGAC TTTGGACCGT TAAGCTGGGC AAAGGAAATG ACAAGGGGAC 1800

GGGGTCTGTG AGAGTCAATT CAGGGGAAAG ATACAAGATT GATTTGTAAA ACCCTTGAAA 1860

TGTAGATTTC TTGTAGATGT ATCCTTCACG TTGTAAATAT GTTTTGTAGA GTGAAGCCAT 1920

GGGAAGCCAT GTGTAACAGA GCTTAGACAT CCAAAACTAA TCAATGCTGA GGTGGCTAAA 1980

TACCTAGCCT TTTACATGTA AACCTGTCTG CAAAATTAGC TTTTTTAAAA AAAAAAAAAA 2040

A 2041

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 187 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Met Arg Gly Asn Ser Arg Asp Ser Val Met Ala Val Thr Thr Gly Gly
1 5 10 15

Gly Thr Arg Ser Leu Thr Arg His Leu Ser Ile His Arg Arg Val Thr
20 25 30

Pro Val Ala Val Ile Thr Asp Met Lys Arg Cys Glu Thr Pro Glu Ile
35 40 45

Leu His His Gln Ile Lys Cys Cys Gly Asp Leu Ile Val Leu Lys Thr
50 55 60

Asn Thr Val Thr Ala Gln Val Thr Val Arg Pro Lys Met Cys Ile Leu
65 70 75 80

```
Thr Glu Leu Glu Arg Gly Met Val Gly Pro Val Thr Leu His Lys Lys
                85                  90                  95

Ile His Thr Thr Thr Val Leu Phe Ile Val Gln Ile His Ile Leu Leu
            100                 105                 110

Ile Gln Ala Ile Thr Gln Ala Lys Leu Gln Met His Leu Met Ile Leu
        115                 120                 125

Gln Met Thr Gly Leu Ser Ile Leu Ala Leu Leu Gly Lys Ser Thr Thr
    130                 135                 140

Thr Ile Val Glu Gln Lys Phe His Asn Gly Lys Asn Gln Lys Ser Gly
145                 150                 155                 160

Leu Lys Glu Asn Arg Asp Lys Lys Lys Gln Thr Arg Trp Gln Ser Thr
                165                 170                 175

Ala Ser Gln Lys Ile Gly Ile Thr Glu Glu Arg
            180                 185
```

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1163 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
GCCCTATCCA CTTAATAGAT GCCAATTCAA AGAGGTTAAA TGATTAGACT AAGGCACCTA     60

ACTTATGTGA GTGTCAGGCT TCAATGCCTG TGTTAGAGCT ACTCCTTCAC ACAAAATAGT    120

TCAGAACATA GAGAAGGACC AAGGTTAATA AATGATTTTC ATCCCAAACA CTAAACATGA    180

TTGATGGGTA GAGGCTGCCC GAAGTACTGT GTAAAGATGG AATCTGAGAT AGAAGAATGC    240

TGTGGTCAAT TAGTAATTCT TGCCCATGGA GGGATTAGTG ACACATGCCT TGTATATTTG    300

TCATCTGTGG CCTAAACTCT GCCCCTGAAG GTTTGTTTTC TAATTCAGAG GTTTAAATTA    360

ATCTAGCCCA CTTAATAAAA CCAGAGATCC TATGGGAAAT TTAGCCTAAG ACAGTGCTGG    420

AAATTGCCAT ATGTTGATAC AAAGAAGTGT TTGGCCACAT TACAGGTCTC AGACTCAACT    480

GCTATGTGTG ACTGCCGCTC TGTGCCTATG TCTTGCTTTT TTGCTGAGTT CCCTATTTCC    540

ATATCTCCAG GTGAATCCAT GAGAAGCGAG AGGGTGGCTG AGAGGCCTGG GCCTCTGGGA    600
```

TTCCACCTTG CTATCTCTGC TCTTCAACCA TTGTTTTAGA CTCTGAACAC CAGATCCTCA 660

TATCTGAAAG TGATTTGGAG ACCTGGGCAT CAAGTGCTCT TTTAAGAAGG GGCTATCCCA 720

GAGGACTGTT CAAAAGTCTC ATTCAATAGA GATGTTGGAG TCCCAGAACA AAGTTAGGGA 780

GCAAACCAGT AACCTATGCT GGTSGTAACA GAGGATCCTA CAATTACGTT TGTTTTTAAG 840

ACAGGATTTT GCTGTGTTGC CCAGACTGGT CTCAAACTCC TGGGTTCAAG AGATCCATCC 900

TCCCACCTCA GTCTCCTGAA AGCTGGGATG ACAGGCACAT GCCACCACAC CTAGCTCCTT 960

ACAACCATTT ATTTTAACTT ATTTCATTTA TAACTGGTAT CTTTCATTTG TATGTGGCAG 1020

CTAGAGATTT ATATAGGATG GAAGTAATTT ATTTTTAATT TAAATATTTC ATGTTGAACT 1080

GTTTGCCTTG TATGGAACAT TTTACTTGGC CAATTCAAAT AAAAATAAAG TCAGCTTTGT 1140

TTGTGACAAA AAAAAAAAAA AAA 1163

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 43 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met Leu Ile Gln Arg Ser Val Trp Pro His Tyr Arg Ser Gln Thr Gln
1 5 10 15

Leu Leu Cys Val Thr Ala Ala Leu Cys Leu Cys Leu Ala Phe Leu Leu
20 25 30

Ser Ser Leu Phe Pro Tyr Leu Gln Val Asn Pro
35 40

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3067 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCGGTGGCTG AGGCGGCTGG GCCTAGGGTG CAGCGGGCGC GTCTGCGGCT GGTGTTGGCG 60

CATCTCTAGA TCCTTTCCCG GAGTTCAGTT ATGGGTGTGA GAGGTTTGCA AGGATTTGTG 120

GGAAGTACCT GCCCACATAT ATGTACAGTA GTAAATTTCA AAGAACTGGC AGAGCACCAC 180

CGAAGCAAGT ATCCTGGATG TACCCCTACC ATTGTGGTTG ATGCCATGTG TTGTCTCAGA 240

TATTGGTATA CTCCAGAATC TTGGATCTGC GGTGGCCAGT GGCGAGAATA CTTTTCTGCT 300

TTGCGAGATT TTGTTAAAAC TTTTACGGCA GCTGGGATCA AGTTGATATT CTTCTTTGAT 360

GGCATGGTGG AGCAGGATAA GAGAGATGAA TGGGTGAAAC GAAGGCTCAA GAACAACAGG 420

GAGATATCCA GGATTTTTCA TTACATCAAG TCACACAAGG AGCAGCCAGG CAGAAATATG 480

TTCTTCATCC CCTCAGGGCT AGCTGTGTTT ACACGATTTG CTCTAAAGAC ACTGGGCCAG 540

GAAACTTTGT GTTCTTTGCA GGAAGCAGAT TATGAGGTAG CTTCCTATGG CCTCCAGCAT 600

AACTGTCTTG GGATTCTGGG GGAAGACACT GATTACCTAA TCTATGACAC TTGTCCCTAC 660

TTTTCAATTA GCGAGCTCTG CCTAGAGAGC CTGGACACCG TCATGCTCTG CAGAGAGAAG 720

CTCTGTGAGA GTCTGGGCCT CTGTGTGGCC GACCTTCCTC TTCTGGCCTG CCTCCTTGGC 780

GACGACATAA TCCCAGAGGG CATGTTTGAA AGCTTTAGGT ACAAATGCTT ATCGTCCTAC 840

ACCTCTGTAA AAGAGAACTT TGACAAAAAA GGTAACATCA TATTAGCTGT GTCAGACCAT 900

ATATCGAAAG TTCTTTACTT GTATCAAGGT GAGAAAAAAT TAGAAGAGAT ATTACCTCTG 960

GGACCAAACA AAGCTCTTTT TTATAAAGGA ATGGCATCAT ATCTTTTACC AGGACAAAAA 1020

TCTCCATGGT TTTTCCAAAA ACCCAAAGGT GTAATAACTT TGGACAAACA AGTAATATCC 1080

ACGAGTTCAG ACGCCGAATC CAGGGAAGAA GTTCCCATGT GTTCAGATGC TGAATCCAGG 1140

CAAGAAGTTC CCATGTGTAC AGGCCCTGAA TCCAGGCGAG AAGTTCCCGT GTATACAGAT 1200

TCTGAACCCA GGCAAGAAGT TCCCATGTGT TCAGACCCTG AACCCAGGCA AGAAGTTCCC 1260

ACATGTACAG GCCCTGAATC CAGGCGAGAA GTTCCCATGT GTTCAGACCC TGAACCCAGG 1320

CAAGAAGTTC CCATGTGTAC AGGCCCTGAA GCCAGGCAAG AAGTTCCCAT GTATACAGAC 1380

TCTGAACCCA GGCAAGAAGT TCCCATGTAT ACAGACTCTG AACCCAGGCA AGAAGTTCCC 1440

ATGTATACAG GCTCTGAACC CAGGCAAGAA GTTCCCATGT ATACAGGCCC TGAATCCAGG 1500

CAAGAAGTTC CCATGTATAC AGGCCCTGAA TCCAGGCAAG AAGTTTTAAT ACGGACAGAC 1560

CCTGAATCTA GGCAAGAAAT TATGTGTACA GGCCATGAAT CCAAACAGGA AGTTCCCATA 1620

TGTACAGATC CTATATCCAA GCAAGAAGAC TCCATGTGTA CACACGCTGA AATCAATCAA 1680

AAATTACCTG TAGCAACAGA TTTTGAATTT AAGCTAGAAG CTCTCATGTG TACAAACCCT 1740

GAAATTAAAC AAGAAGACCC CACAAATGTG GGGCCTGAAG TAAAGCAACA AGTAACCATG 1800

GTTTCAGACA CTGAAATCTT AAAGGTTGCT AGAACACATC ACGTCCAAGC AGAAAGCTAC 1860

CTGGTGTACA ACATCATGAG CAGTGGAGAG ATTGAATGCA GCAACACCCT AGAAGATGAG 1920

CTTGACCAGG CCTTACCCAG CCAGGCCTTC ATTTACCGTC CCATTCGACA GCGGGTCTAC 1980

TCACTCTTAC TGGAGGACTG TCAAGATGTC ACCAGCACCT GCCTAGCTGT CAAGGAGTGG 2040

TTTGTGTATC CTGGGAACCC ACTGAGGCAC CCGGACCTCG TCAGGCCGCT GCAGATGACC 2100

ATTCCAGGGG GAACGCCTAG TTTGAAAATA TTATGGCTGA ACCAAGAGCC AGAAATACAG 2160

GTTCGGCGCT TGGACACACT CCTAGCCTGT TTCAATCTTT CCTCCTCAAG AGAAGAGCTG 2220

CAGGCTGTCG AAAGCCCATT TCAAGCTTTG TGCTGCCTCT TGATCTACCT CTTTGTCCAG 2280

GTGGACACGC TTTGCCTGGA GGATTTGCAT GCGTTTATTG CGCAGGCCTT GTGCCTCCAA 2340

GGAAAATCCA CCTCGCAGCT TGTAAATCTA CAGCCTGATT ACATCAACCC CAGAGCCGTG 2400

CAGCTGGGCT CCCTTCTCGT CCGCGGCCTC ACCACTCTGG TTTTAGTCAA CAGCGCATGT 2460

GGCTTCCCCT GGAAGACGAG TGATTTCATG CCCTGGAATG TATTTGACGG GAAGCTTTTT 2520

CATCAGAAGT ACTTGCAATC TGAAAAGGGT TATGCTGTGG AGGTTCTTTT AGAACAAAAT 2580

GGAGGTGGGG AAGACAGGGC TCCAGCTACC ACAGGACGGG CTCTGGGTAT AGCCGTTCCA 2640

GTCAGGGACA GCCGTGGAGA GACCAGGGAC CAGGAAGCAG ACAGTATGAG CATGACCAGT 2700

GGAGAAGGTA CTAGTCAACC TCCAGAAAGA GTATGGAGAG AAAAAGAGGC ACACCTGGAC 2760

GCAGAGCCCT GCCAGCGCCC TCCTCTGCTG TTGCAGCTGC AAGGAGACCA TGCCTGTGGG 2820

AGCCAGGCCT CGCTTGCATG AAGAAGGAAC GATGCCTTTT TCAATGGTGT CTCCCTCCCA 2880

TTGTGCAGAA GAGCTTTTGT TGGCTTCTCT CCCGAGCTTG TGCCTGATTC TGTGGCCCAA 2940

AACAATCATT GTTAACATCT TCATGTGTTT CATTCTGATC TTTCATTCAT ATATATGATG 3000

CCTAGCTAAT TTCATTTTAA AATAAATGGG AATCTGTTGT AAAAAAAAAA AAAAAAAAAA 3060

AAAAAAA 3067

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 916 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
Met Gly Val Arg Gly Leu Gln Gly Phe Val Gly Ser Thr Cys Pro His
1               5                  10                  15

Ile Cys Thr Val Val Asn Phe Lys Glu Leu Ala Glu His His Arg Ser
            20                  25                  30

Lys Tyr Pro Gly Cys Thr Pro Thr Ile Val Val Asp Ala Met Cys Cys
        35                  40                  45

Leu Arg Tyr Trp Tyr Thr Pro Glu Ser Trp Ile Cys Gly Gly Gln Trp
    50                  55                  60

Arg Glu Tyr Phe Ser Ala Leu Arg Asp Phe Val Lys Thr Phe Thr Ala
65                  70                  75                  80

Ala Gly Ile Lys Leu Ile Phe Phe Phe Asp Gly Met Val Glu Gln Asp
                85                  90                  95

Lys Arg Asp Glu Trp Val Lys Arg Arg Leu Lys Asn Asn Arg Glu Ile
            100                 105                 110

Ser Arg Ile Phe His Tyr Ile Lys Ser His Lys Glu Gln Pro Gly Arg
        115                 120                 125

Asn Met Phe Phe Ile Pro Ser Gly Leu Ala Val Phe Thr Arg Phe Ala
    130                 135                 140

Leu Lys Thr Leu Gly Gln Glu Thr Leu Cys Ser Leu Gln Glu Ala Asp
145                 150                 155                 160

Tyr Glu Val Ala Ser Tyr Gly Leu Gln His Asn Cys Leu Gly Ile Leu
                165                 170                 175

Gly Glu Asp Thr Asp Tyr Leu Ile Tyr Asp Thr Cys Pro Tyr Phe Ser
            180                 185                 190

Ile Ser Glu Leu Cys Leu Glu Ser Leu Asp Thr Val Met Leu Cys Arg
        195                 200                 205

Glu Lys Leu Cys Glu Ser Leu Gly Leu Cys Val Ala Asp Leu Pro Leu
    210                 215                 220

Leu Ala Cys Leu Leu Gly Asp Asp Ile Ile Pro Glu Gly Met Phe Glu
```

225 230 235 240

Ser Phe Arg Tyr Lys Cys Leu Ser Ser Tyr Thr Ser Val Lys Glu Asn
245 250 255

Phe Asp Lys Lys Gly Asn Ile Ile Leu Ala Val Ser Asp His Ile Ser
260 265 270

Lys Val Leu Tyr Leu Tyr Gln Gly Glu Lys Lys Leu Glu Glu Ile Leu
275 280 285

Pro Leu Gly Pro Asn Lys Ala Leu Phe Tyr Lys Gly Met Ala Ser Tyr
290 295 300

Leu Leu Pro Gly Gln Lys Ser Pro Trp Phe Phe Gln Lys Pro Lys Gly
305 310 315 320

Val Ile Thr Leu Asp Lys Gln Val Ile Ser Thr Ser Ser Asp Ala Glu
325 330 335

Ser Arg Glu Glu Val Pro Met Cys Ser Asp Ala Glu Ser Arg Gln Glu
340 345 350

Val Pro Met Cys Thr Gly Pro Glu Ser Arg Arg Glu Val Pro Val Tyr
355 360 365

Thr Asp Ser Glu Pro Arg Gln Glu Val Pro Met Cys Ser Asp Pro Glu
370 375 380

Pro Arg Gln Glu Val Pro Thr Cys Thr Gly Pro Glu Ser Arg Arg Glu
385 390 395 400

Val Pro Met Cys Ser Asp Pro Glu Pro Arg Gln Glu Val Pro Met Cys
405 410 415

Thr Gly Pro Glu Ala Arg Gln Glu Val Pro Met Tyr Thr Asp Ser Glu
420 425 430

Pro Arg Gln Glu Val Pro Met Tyr Thr Asp Ser Glu Pro Arg Gln Glu
435 440 445

Val Pro Met Tyr Thr Gly Ser Glu Pro Arg Gln Glu Val Pro Met Tyr
450 455 460

Thr Gly Pro Glu Ser Arg Gln Glu Val Pro Met Tyr Thr Gly Pro Glu
465 470 475 480

Ser Arg Gln Glu Val Leu Ile Arg Thr Asp Pro Glu Ser Arg Gln Glu
485 490 495

Ile Met Cys Thr Gly His Glu Ser Lys Gln Glu Val Pro Ile Cys Thr
500 505 510

Asp Pro Ile Ser Lys Gln Glu Asp Ser Met Cys Thr His Ala Glu Ile
515 520 525

```
Asn Gln Lys Leu Pro Val Ala Thr Asp Phe Glu Phe Lys Leu Glu Ala
    530                 535                 540

Leu Met Cys Thr Asn Pro Glu Ile Lys Gln Glu Asp Pro Thr Asn Val
545                 550                 555                 560

Gly Pro Glu Val Lys Gln Gln Val Thr Met Val Ser Asp Thr Glu Ile
                565                 570                 575

Leu Lys Val Ala Arg Thr His His Val Gln Ala Glu Ser Tyr Leu Val
            580                 585                 590

Tyr Asn Ile Met Ser Ser Gly Glu Ile Glu Cys Ser Asn Thr Leu Glu
        595                 600                 605

Asp Glu Leu Asp Gln Ala Leu Pro Ser Gln Ala Phe Ile Tyr Arg Pro
    610                 615                 620

Ile Arg Gln Arg Val Tyr Ser Leu Leu Leu Glu Asp Cys Gln Asp Val
625                 630                 635                 640

Thr Ser Thr Cys Leu Ala Val Lys Glu Trp Phe Val Tyr Pro Gly Asn
                645                 650                 655

Pro Leu Arg His Pro Asp Leu Val Arg Pro Leu Gln Met Thr Ile Pro
            660                 665                 670

Gly Gly Thr Pro Ser Leu Lys Ile Leu Trp Leu Asn Gln Glu Pro Glu
        675                 680                 685

Ile Gln Val Arg Arg Leu Asp Thr Leu Leu Ala Cys Phe Asn Leu Ser
    690                 695                 700

Ser Ser Arg Glu Glu Leu Gln Ala Val Glu Ser Pro Phe Gln Ala Leu
705                 710                 715                 720

Cys Cys Leu Leu Ile Tyr Leu Phe Val Gln Val Asp Thr Leu Cys Leu
                725                 730                 735

Glu Asp Leu His Ala Phe Ile Ala Gln Ala Leu Cys Leu Gln Gly Lys
            740                 745                 750

Ser Thr Ser Gln Leu Val Asn Leu Gln Pro Asp Tyr Ile Asn Pro Arg
        755                 760                 765

Ala Val Gln Leu Gly Ser Leu Leu Val Arg Gly Leu Thr Thr Leu Val
    770                 775                 780

Leu Val Asn Ser Ala Cys Gly Phe Pro Trp Lys Thr Ser Asp Phe Met
785                 790                 795                 800

Pro Trp Asn Val Phe Asp Gly Lys Leu Phe His Gln Lys Tyr Leu Gln
                805                 810                 815

Ser Glu Lys Gly Tyr Ala Val Glu Val Leu Leu Glu Gln Asn Gly Gly
```

```
                820                 825                 830

Gly Glu Asp Arg Ala Pro Ala Thr Thr Gly Arg Ala Leu Gly Ile Ala
        835                 840                 845

Val Pro Val Arg Asp Ser Arg Gly Glu Thr Arg Asp Gln Glu Ala Asp
    850                 855                 860

Ser Met Ser Met Thr Ser Gly Glu Gly Thr Ser Gln Pro Pro Glu Arg
865                 870                 875                 880

Val Trp Arg Glu Lys Glu Ala His Leu Asp Ala Glu Pro Cys Gln Arg
                885                 890                 895

Pro Pro Leu Leu Leu Gln Leu Gln Gly Asp His Ala Cys Gly Ser Gln
            900                 905                 910

Ala Ser Leu Ala
        915
```

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1914 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

```
AGCTGTCTGC TCTCCTGGCA GGAATCGCTG AGGGAGGGAA ACGCGGCTCT GAATCAGCCC      60

AGAACGAGCC TTCGGGAAGC TCACCCTCCG ATCTCGGTGT GATTGTTGTG ATTGTTGTGA     120

TTTCCTGTCT CGTTTGCCTT GACCGCCATG TGAAAGAATC TGTTCCCCAG CTAGGTGGGG     180

AAAATTCACA GGTGGGCTGT CTGTAGAGAG AACTGGCTGA TTAAAGGCTT CTCGTCCCGA     240

TTTTGTGATA GCCAAGTGCT TGGCCTGGTC GACGGTCTTT GCTCCTTTAC AAATAAAGTG     300

TTCTGTTTCA GTTCGTCCCA AGTTTTCCAT GAAGGGCAGT GGTTCCCTGA CCTCCCAGGT     360

GCCTGGGCTT CCCCAGGTTC CTGATCTGGG GCTTGGGGCC CTGTGTTTGG GGATCGTGGC     420

ACTGTGTGCA CCAGCCTGGA AGCACTGGGC CAGTCTTGGC CAAGCTTTCC ATCAGGGATG     480

ATTTGATCTT GGTGCTACAG GTCTGTGGTA CGACCATTGT TCCACACCAC ATGTCATTAA     540

TAATGCTTCC CATGCTTCTG CTTGCAAATG ACCAGCCTTC CAAACAGCCA GAGCTGTTTC     600
```

GAGGTGTTTC TGCAGGCAGG TGCAGGCGTG CCCTCAAATA AGCTTTGCCA ATGGAGTCTC 660

AGCAAGAGCA AAACCTGGTC AGGAAAGACA AAGCCTGGGA ATCCACCCCC ATGCCCTGCA 720

GGTTGGCTGG CCCTGGAGCC ATTTATTATA GTGCTAATCA TGTTTCTAGG CAGGTGCAGA 780

TGGCAAGGGC AGTGTCTTGG TGAGCTTTTT AGCACGAAGA GCCAGGTCTG TCGAAGCCTT 840

TGTGAGAGCT GGAAACGCAG GTGTGCTGGG CATGCGCAGT ATGGGGTTTC GGGCTCAGGG 900

CTTGCCCTTT GGCATCAGAC AGACCTGGCT TCGCATCCTG GATTTGCTTC TGACGTGCAC 960

CCTTCCCTTT GGGTCTCGTG ATGTGAAATG GAGATGTTGT CATTTGTGAG GGCTCCATGA 1020

AGTTTCGTTG AAATGACAAA TACTAATTTC TTCATCTGTG AAATGGAGAT AATAGTGCTG 1080

ACCTCAGAAC AGCTGAGAGG ACTAAATGAA ATGATGTTGG ATGTAGCCAT AAAGAACGAA 1140

GTCAGGCACT GGTGCACGCC TGGAATCCCA GCTCTTGGGA GACCGAGACA GGTGGATTGC 1200

TTGAGCTCAG GAGTTTGAGA CCAGCCTGAG CAACATAGGG AGGTCCAGTC TCTACAAAAA 1260

ATATGAAAAG TAGCTGGGCG TGGTGGCGCA TGCCTGTAGT CCCACTACTT GGAAGGCTTC 1320

GTTGGGAGGA TCACTTGAGC CCAGAAGATT GAGGCTGCAG TAAGCCGTGA TCGTGCCACT 1380

GCATTCCAGC CTGGGCAACA GAGCGAGACA CTGTCTCAAA TAAAAAAGAT GGGAATAGTA 1440

GACACTGGGG GCTCCAGAAG GAGGGAGGGA GGGAGGAAGG GGAGGAAGGG CTGAAATGCT 1500

TTCTATTGGA TACTATCTGG GCATATTACT TCCTGTGGTT CACTGTCTGG GTGACAGGAT 1560

TCATAGAAGC CCAAACTTTA GCACCACGCA GCATACCCTT GTAACAAAGC CGCACACGTA 1620

CGCCCTCAAG CTAAAACAAA AGTGGACCGG GAGGCCGAGG TCGGGGGATC ATGAGGTCAG 1680

GAGTTTGAGA CCAGCCTGGC AGATAACGGT GAAACCCCGT CTCTACTAAA AATACCAAAA 1740

AAAGTTAGCC GGACATGGTG GCAGGTGCCT GTAGTCCCAG CTACTTGGGA GGCTGGGGCA 1800

GAAGAATCGC TTGAACCCAG GAGGCGGAGG TTGCAGTGAG CCGAGATTGC GCCACTGCAC 1860

TCCAGCCTGT GCGACAGAGT GAGACTCCGT CTCAAAAAAA AAAAAAAAAA AAAA 1914

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 137 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

```
Met Thr Ser Leu Pro Asn Ser Gln Ser Cys Phe Glu Val Phe Leu Gln
1               5                  10                  15

Ala Gly Ala Gly Val Pro Ser Asn Lys Leu Cys Gln Trp Ser Leu Ser
            20                  25                  30

Lys Ser Lys Thr Trp Ser Gly Lys Thr Lys Pro Gly Asn Pro Pro Pro
        35                  40                  45

Cys Pro Ala Gly Trp Leu Ala Leu Glu Pro Phe Ile Ile Val Leu Ile
    50                  55                  60

Met Phe Leu Gly Arg Cys Arg Trp Gln Gly Gln Cys Leu Gly Glu Leu
65                  70                  75                  80

Phe Ser Thr Lys Ser Gln Val Cys Arg Ser Leu Cys Glu Ser Trp Lys
                85                  90                  95

Arg Arg Cys Ala Gly His Ala Gln Tyr Gly Val Ser Gly Ser Gly Leu
            100                 105                 110

Ala Leu Trp His Gln Thr Asp Leu Ala Ser His Pro Gly Phe Ala Ser
        115                 120                 125

Asp Val His Pro Ser Leu Trp Val Ser
    130                 135
```

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 575 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

```
CCGACTCCCT TCTTTATGGC GTCGCTCCTG TGCTGTGGGC CGAAGCTGGC CGCCTGCGGC     60

ATCGTCCTCA GCGCCTGGGG AGTGATCATG TTGATAATGC TCGGAATATT TTTCAATGTC     120

CATTCCGCTG TGTTGATTGA GGACGTTCCC TTCACGGAGA AAGATTTTGA GAATGGCCCC     180

CAGAACATAT ACAACCTTTA CGAGCAAGTC AGCTACAACT GTTTCATCGC TGCAGGCCTT     240

TACCTCCTCC TCGGAGGCTT CTCTTTCTGC CAAGTTCGGC TCAATAAGCG CAAGGAATAC     300
```

ATGGTGCGCT AGGGCCCCGG CGCGTTTCCC CGCTCCAGCC CCTCCTCTAT TTAAAGACTC 360

CCTGCACCGT GTCACCCAGG TCGCGTCCCA CCCTTGCCGG CGCCCTCTGT GGGACTGGGT 420

TTCCCGGGCG AGAGACTGAA TCCCTTCTCC CATCTCTGGC ATCCGGCCCC CGTGGAGAGG 480

GCTGAGGCTG GGGGGCTGTT CCGTCTCTCC ACCCTTCGCT GTGTCCCGTA TCTCAATAAA 540

GAGAATCTGC TCTCTTCAAA AAAAAAAAAA AAAAA 575

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 98 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Ala Ser Leu Leu Cys Cys Gly Pro Lys Leu Ala Ala Cys Gly Ile
1 5 10 15

Val Leu Ser Ala Trp Gly Val Ile Met Leu Ile Met Leu Gly Ile Phe
20 25 30

Phe Asn Val His Ser Ala Val Leu Ile Glu Asp Val Pro Phe Thr Glu
35 40 45

Lys Asp Phe Glu Asn Gly Pro Gln Asn Ile Tyr Asn Leu Tyr Glu Gln
50 55 60

Val Ser Tyr Asn Cys Phe Ile Ala Ala Gly Leu Tyr Leu Leu Leu Gly
65 70 75 80

Gly Phe Ser Phe Cys Gln Val Arg Leu Asn Lys Arg Lys Glu Tyr Met
85 90 95

Val Arg

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GNAGCCCAGGA GTCTTCTCAA CCTCTTCC 29

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ANCAGTCGCAA GTGCATAGTA ACCCAGTA 29

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TNCTCAGCTTT TATTTGGTTC TGAGTGTT 29

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TNTGCTCAGAC CAGTCATCTG CAGAATCA 29

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

TNCAGCACTGT CTTAGGCTAA ATTTCCCA 29

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

GNATTCGGCGT CTGAACTCGT GGATATTA 29

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

ANATGCCCAGA TAGTATCCAA TAGAAAGC 29

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 29 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid
(A) DESCRIPTION: /desc = "oligonucleotide"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

CNACAGCACAG GAGCGACGCC ATAAAGAA 29

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 543 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS:
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

```
Met Val Met Tyr Ala Arg Lys Gln Gln Arg Leu Ser Asp Gly Cys His
1               5                  10                  15

Asp Arg Arg Gly Asp Ser Gln Pro Tyr Gln Ala Leu Lys Tyr Ser Ser
            20                  25                  30

Lys Ser His Pro Ser Ser Gly Asp His Arg His Glu Lys Met Arg Asp
        35                  40                  45

Ala Gly Asp Pro Ser Pro Pro Asn Lys Met Leu Arg Arg Ser Asp Ser
    50                  55                  60

Pro Glu Asn Lys Tyr Ser Asp Ser Thr Gly His Ser Lys Ala Lys Asn
65                  70                  75                  80

Val His Thr His Arg Val Arg Glu Arg Asp Gly Gly Thr Ser Tyr Ser
                85                  90                  95
```

```
Pro Gln Glu Asn Ser His Asn His Ser Ala Leu His Ser Ser Asn Ser
            100                 105                 110

His Ser Ser Asn Pro Ser Asn Asn Pro Ser Lys Thr Ser Asp Ala Pro
        115                 120                 125

Tyr Asp Ser Ala Asp Asp Trp Ser Glu His Ile Ser Ser Ser Gly Lys
    130                 135                 140

Lys Tyr Tyr Tyr Asn Cys Arg Thr Glu Val Ser Gln Trp Glu Lys Pro
145                 150                 155                 160

Lys Glu Trp Leu Glu Arg Glu Gln Arg Gln Lys Glu Ala Asn Lys Met
                165                 170                 175

Ala Val Asn Ser Phe Pro Lys Asp Arg Asp Tyr Arg Arg Glu Val Met
            180                 185                 190

Gln Ala Thr Ala Thr Ser Gly Phe Ala Ser Gly Lys Ser Thr Ser Gly
        195                 200                 205

Asp Lys Pro Val Ser His Ser Cys Thr Thr Pro Ser Thr Ser Ser Ala
    210                 215                 220

Ser Gly Leu Asn Pro Thr Ser Ala Pro Pro Thr Ser Ala Ser Ala Val
225                 230                 235                 240

Pro Val Ser Pro Val Pro Gln Ser Pro Ile Pro Pro Leu Leu Gln Asp
                245                 250                 255

Pro Asn Leu Leu Arg Gln Leu Leu Pro Ala Leu Gln Ala Thr Leu Gln
            260                 265                 270

Leu Asn Asn Ser Asn Val Asp Ile Ser Lys Ile Asn Glu Val Leu Thr
        275                 280                 285

Ala Ala Val Thr Gln Ala Ser Leu Gln Ser Ile Ile His Lys Phe Leu
    290                 295                 300

Thr Ala Gly Pro Ser Ala Phe Asn Ile Thr Ser Leu Ile Ser Gln Ala
305                 310                 315                 320

Ala Gln Leu Ser Thr Gln Ala Gln Pro Ser Asn Gln Ser Pro Met Ser
                325                 330                 335

Leu Thr Ser Asp Ala Ser Ser Pro Arg Ser Tyr Val Ser Pro Arg Ile
            340                 345                 350

Ser Thr Pro Gln Thr Asn Thr Val Pro Ile Lys Pro Leu Ile Ser Thr
        355                 360                 365

Pro Pro Val Ser Ser Gln Pro Lys Val Ser Thr Pro Val Val Lys Gln
    370                 375                 380

Gly Pro Val Ser Gln Ser Ala Thr Gln Gln Pro Val Thr Ala Asp Lys
```

385 390 395 400

Gln Gln Gly His Glu Pro Val Ser Pro Arg Ser Leu Gln Arg Ser Ser
405 410 415

Gln Arg Ser Pro Ser Pro Gly Pro Asn His Thr Ser Asn Ser Ser Asn
420 425 430

Ala Ser Asn Ala Thr Val Val Pro Gln Asn Ser Ser Ala Arg Ser Thr
435 440 445

Cys Ser Leu Thr Pro Ala Leu Ala Ala His Phe Ser Glu Asn Leu Ile
450 455 460

Lys His Val Gln Gly Trp Pro Ala Asp His Ala Glu Lys Gln Ala Ser
465 470 475 480

Arg Leu Arg Glu Glu Ala His Asn Met Gly Thr Ile His Met Ser Glu
485 490 495

Ile Cys Thr Glu Leu Lys Asn Leu Arg Ser Leu Val Arg Val Cys Glu
500 505 510

Ile Gln Ala Thr Leu Arg Glu Gln Arg Ile Leu Phe Leu Arg Gln Gln
515 520 525

Ile Lys Glu Leu Glu Lys Leu Lys Asn Gln Asn Ser Phe Met Val
530 535 540

What is claimed is:

1. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 99 to nucleotide 902;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 162 to nucleotide 902;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:1 from nucleotide 87 to nucleotide 219;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone ci25_4 deposited under accession number ATCC 98415;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone ci25_4 deposited under accession number ATCC 98415;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone ci25_4 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone ci25_4 deposited under accession number ATCC 98415;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:2;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:2 having biological activity, the fragment comprising the amino acid sequence from amino acid 129 to amino acid 138 of SEQ ID NO:2;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

2. The polynucleotide of claim 1 wherein said polynucleotide is operably linked to at least one expression control sequence.

3. A host cell transformed with the polynucleotide of claim 2.

4. The host cell of claim 3, wherein said cell is a mammalian cell.

5. A process for producing a protein encoded by the polynucleotide of claim 2, which process comprises:

(a) growing a culture of the host cell of claim 3 in a suitable culture medium; and

(b) purifying said protein from the culture.

6. A protein produced according to the process of claim 5.

7. The protein of claim 6 comprising a mature protein.

8. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:2;

(b) fragments of the amino acid sequence of SEQ ID NO:2 comprising the amino acid sequence from amino acid 129 to amino acid 138 of SEQ ID NO:2; and

(c) the amino acid sequence encoded by the cDNA insert of clone ci25_4 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins.

9. The protein of claim 8, wherein said protein comprises the amino acid sequence of SEQ ID NO:2.

10. A composition comprising the protein of claim 8 and a pharmaceutically acceptable carrier.

11. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:1.

12. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 283 to nucleotide 1158;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:3 from nucleotide 1 to nucleotide 789;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone da228_6 deposited under accession number ATCC 98415;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone da228_6 deposited under accession number ATCC 98415;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone da228_6 deposited under accession number ATCC 98415;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone da228_6 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:4;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:4 having biological activity, the fragment comprising the amino acid sequence from amino acid 141 to amino acid 150 of SEQ ID NO:4;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

13. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:4;

(b) the amino acid sequence of SEQ ID NO:4 from amino acid 1 to amino acid 169;

(c) fragments of the amino acid sequence of SEQ ID NO:4 comprising the amino acid sequence from amino acid 141 to amino acid 150 of SEQ ID NO:4; and

(d) the amino acid sequence encoded by the cDNA insert of clone da228_6 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins.

14. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:3.

15. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 152 to nucleotide 2182;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:5 from nucleotide 2 to nucleotide 931;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone du410_5 deposited under accession number ATCC 98415;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone du410_5 deposited under accession number ATCC 98415;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone du410_5 deposited under accession number ATCC 98415;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone du410_5 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:6;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:6 having biological activity, the fragment comprising the amino acid sequence from amino acid 333 to amino acid 342 of SEQ ID NO:6;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

16. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:6;

(b) the amino acid sequence of SEQ ID NO:6 from amino acid 1 to amino acid 260;

(c) fragments of the amino acid sequence of SEQ ID NO:6 comprising the amino acid sequence from amino acid 333 to amino acid 342 of SEQ ID NO:6; and

(d) the amino acid sequence encoded by the cDNA insert of clone du410_5 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins.

17. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:5.

18. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 51 to nucleotide 611;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:7 from nucleotide 1 to nucleotide 525;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone eh80_1 deposited under accession number ATCC 98415;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone eh80_1 deposited under accession number ATCC 98415;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone eh80_1 deposited under accession number ATCC 98415;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone eh80_1 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:8;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:8 having biological activity, the fragment comprising the amino acid sequence from amino acid 88 to amino acid 97 of SEQ ID NO:8;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

19. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:8;

(b) the amino acid sequence of SEQ ID NO:8 from amino acid 1 to amino acid 158;

(c) fragments of the amino acid sequence of SEQ ID NO:8 comprising the amino acid sequence from amino acid 88 to amino acid 97 of SEQ ID NO:8; and

(d) the amino acid sequence encoded by the cDNA insert of clone eh80_1 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins.

20. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:7.

21. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 431 to nucleotide 559;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 518 to nucleotide 559;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:9 from nucleotide 190 to nucleotide 547;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone er369_1 deposited under accession number ATCC 98415;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone er369_1 deposited under accession number ATCC 98415;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone er369_1 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone er369_1 deposited under accession number ATCC 98415;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:10;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:10 having biological activity, the fragment comprising the amino acid sequence from amino acid 16 to amino acid 25 of SEQ ID NO:10;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

22. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:10;

(b) the amino acid sequence of SEQ ID NO:10 from amino acid 1 to amino acid 39;

(c) fragments of the amino acid sequence of SEQ ID NO:10 comprising the amino acid sequence from amino acid 16 to amino acid 25 of SEQ ID NO:10; and

(d) the amino acid sequence encoded by the cDNA insert of clone er369_1 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins.

23. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:9.

24. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 91 to nucleotide 2838;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 2209 to nucleotide 2838;

(d) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:11 from nucleotide 839 to nucleotide 1197;

(e) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fh123_5 deposited under accession number ATCC 98415;

(f) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fh123_5 deposited under accession number ATCC 98415;

(g) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fh123_5 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fh123_5 deposited under accession number ATCC 98415;

(i) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:12;

(j) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:12 having biological activity, the fragment comprising the amino acid sequence from amino acid 453 to amino acid 462 of SEQ ID NO:12;

(k) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(h) above;

(l) a polynucleotide which encodes a species homologue of the protein of (i) or (j) above ; and

(m) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(j).

25. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:12;

(b) the amino acid sequence of SEQ ID NO:12 from amino acid 251 to amino acid 369;

(c) fragments of the amino acid sequence of SEQ ID NO:12 comprising the amino acid sequence from amino acid 453 to amino acid 462 of SEQ ID NO:12; and

(d) the amino acid sequence encoded by the cDNA insert of clone fh123_5 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins.

26. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:11.

27. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 568 to nucleotide 978;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:13 from nucleotide 1084 to nucleotide 1854;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fm60_1 deposited under accession number ATCC 98415;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fm60_1 deposited under accession number ATCC 98415;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fm60_1 deposited under accession number ATCC 98415;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fm60_1 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:14;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:14 having biological activity, the fragment comprising the amino acid sequence from amino acid 63 to amino acid 72 of SEQ ID NO:14;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

28. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:14;

(b) fragments of the amino acid sequence of SEQ ID NO:14 comprising the amino acid sequence from amino acid 63 to amino acid 72 of SEQ ID NO:14; and

(c) the amino acid sequence encoded by the cDNA insert of clone fm60_1 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins.

29. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:13.

30. An isolated polynucleotide selected from the group consisting of:

(a) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15;

(b) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 16 to nucleotide 309;

(c) a polynucleotide comprising the nucleotide sequence of SEQ ID NO:15 from nucleotide 127 to nucleotide 309;

(d) a polynucleotide comprising the nucleotide sequence of the full-length protein coding sequence of clone fr473_2 deposited under accession number ATCC 98415;

(e) a polynucleotide encoding the full-length protein encoded by the cDNA insert of clone fr473_2 deposited under accession number ATCC 98415;

(f) a polynucleotide comprising the nucleotide sequence of a mature protein coding sequence of clone fr473_2 deposited under accession number ATCC 98415;

(g) a polynucleotide encoding a mature protein encoded by the cDNA insert of clone fr473_2 deposited under accession number ATCC 98415;

(h) a polynucleotide encoding a protein comprising the amino acid sequence of SEQ ID NO:16;

(i) a polynucleotide encoding a protein comprising a fragment of the amino acid sequence of SEQ ID NO:16 having biological activity, the fragment comprising the amino acid sequence from amino acid 44 to amino acid 53 of SEQ ID NO:16;

(j) a polynucleotide which is an allelic variant of a polynucleotide of (a)-(g) above;

(k) a polynucleotide which encodes a species homologue of the protein of (h) or (i) above ; and

(l) a polynucleotide that hybridizes under stringent conditions to any one of the polynucleotides specified in (a)-(i).

31. A protein comprising an amino acid sequence selected from the group consisting of:

(a) the amino acid sequence of SEQ ID NO:16;

(b) the amino acid sequence of SEQ ID NO:16 from amino acid 1 to amino acid 58;

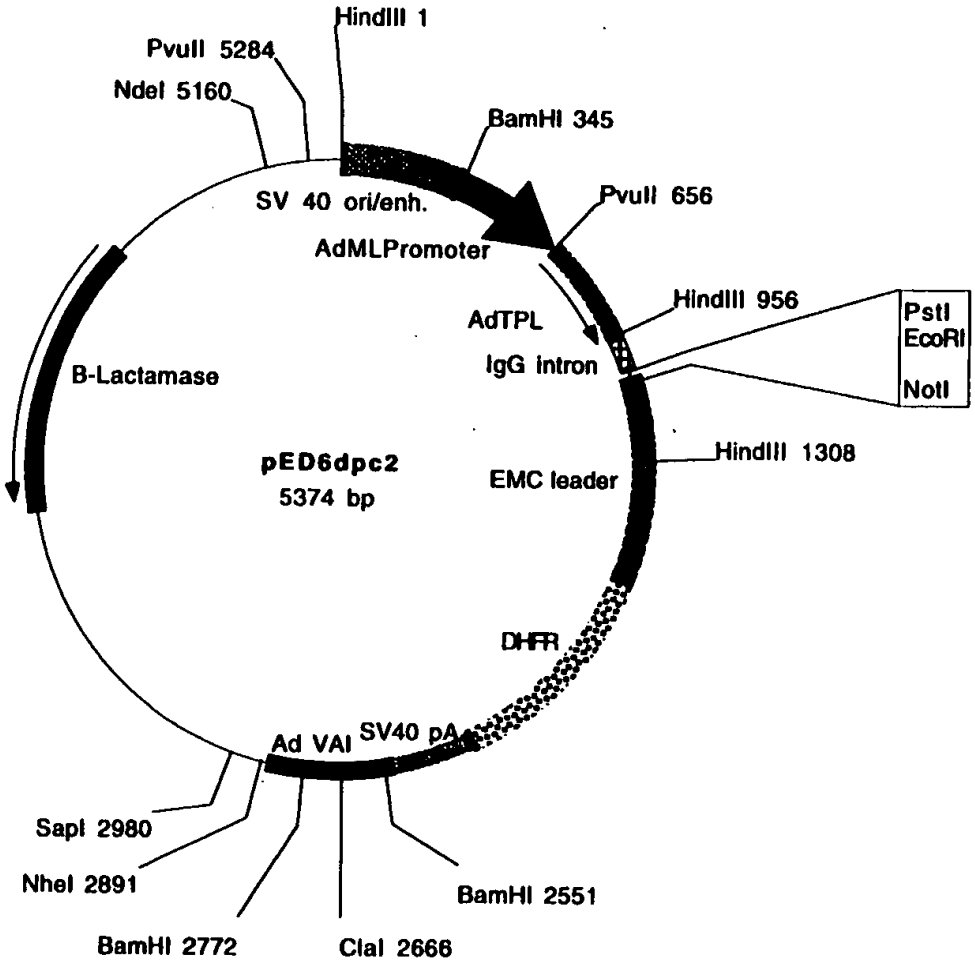
(c) fragments of the amino acid sequence of SEQ ID NO:16 comprising the amino acid sequence from amino acid 44 to amino acid 53 of SEQ ID NO:16; and

(d) the amino acid sequence encoded by the cDNA insert of clone fr473_2 deposited under accession number ATCC 98415;

the protein being substantially free from other mammalian proteins.

32. An isolated gene corresponding to the cDNA sequence of SEQ ID NO:15.

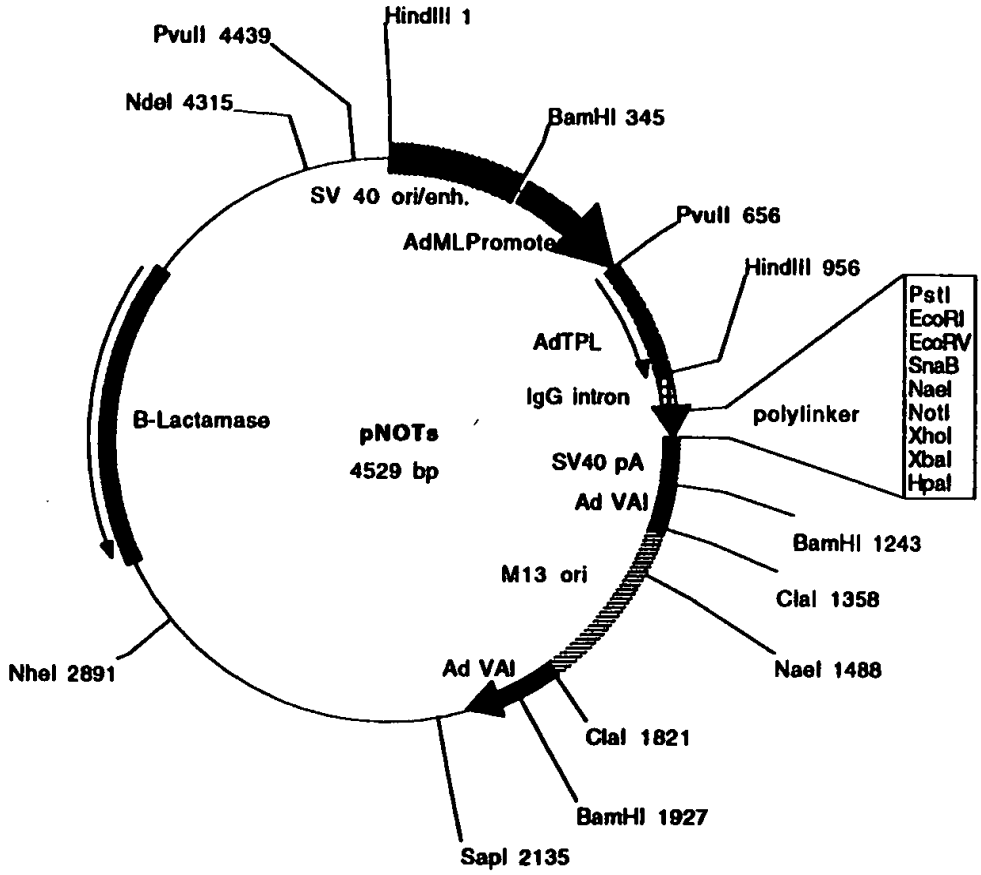
FIGURE 1A



Plasmid name: pED6dpc2
Plasmid size: 5374 bp

Comments/References: pED6dpc2 is derived from pED6dpc1 by insertion of a new polylinker to facilitate cDNA cloning. SST cDNAs are cloned between EcoRI and NotI. pED vectors are described in Kaufman et al.(1991), NAR 19: 4485-4490.

FIGURE 1B



Plasmid name: pNOTs
Plasmid size: 4529 bp

Comments/References: pNOTs is a derivative of pMT2 (Kaufman et al,1989. Mol.Cell.Biol.9:1741-1750). DHFR was deleted and a new polylinker was inserted between EcoRI and HpaI. M13 origin of replication was inserted in the ClaI site. SST cDNAs are cloned between EcoRI and NotI

INTERNATIONAL SEARCH REPORT

Interna[illegible]al Application No
PCT/US 98/08336

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/47 A61K38/17

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 383 129 A (MOLECULAR DIAGNOSTICS INC.) 22 August 1990 see page 9, line 35-44 see page 9 - page 10; claims see figures 8,11,13	1-11
A	WO 97 07198 A (GENETICS INSTITUTE INC.) 27 February 1997	
A	US 5 536 637 A (JACOBS KENNETH) 16 July 1996 cited in the application	

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
8 July 1998	08.10.98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,	Macchia, G

INTERNATIONAL SEARCH REPORT

I. ntional application No.
PCT/US 98/ 08336

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see further information sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

1-11

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1-11

   Polynucleotide comprising the nucleotide sequence of Seq.ID:1 and encoding a polypeptide of Seq.ID:2 or fragments. Polynucleotide fragments, variants, homologues, gene thereof. Host cell transformed with said polynucleotide. Protein comprising an amino acid sequence of Seq.ID:2 or fragments thereof. Process for producing said protein.

2. Claims: 12-14

   Polynucleotide comprising the nucleotide sequence of Seq.ID:3 and encoding a polypeptide of Seq.ID:4 or fragments. Polynucleotide fragments, variants, homologues, gene thereof. Protein comprising an amino acid sequence of Seq.ID:4 or fragments thereof.

3. Claims: 15-17

   As invention 2 but concerning Seq.ID:5 and 6.

4. Claims: 18-20

   As invention 2 but concerning Seq.ID:7 and 8.

5. Claims: 21-23

   As invention 2 but concerning Seq.ID:9 and 10.

6. Claims: 24-26

   As invention 2 but concerning Seq.ID:11 and 12.

7. Claims: 27-29

   As invention 2 but concerning Seq.ID:13 and 14.

8. Claims: 30-32

   As invention 2 but concerning Seq.ID:15 and 16.

INTERNATIONAL SEARCH REPORT

.mation on patent family members

Inter nal Application No
PCT/US 98/08336

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0383129 A	22-08-1990	AT 122394 T	15-05-1995
		AU 626636 B	06-08-1992
		AU 4988390 A	23-08-1990
		CA 2010275 A	17-08-1990
		DE 69019190 D	14-06-1995
		DE 69019190 T	14-09-1995
		DK 383129 T	09-10-1995
		ES 2072928 T	01-08-1995
		GR 3017019 T	30-11-1995
		IE 66997 B	21-02-1996
		IL 93397 A	10-06-1997
		JP 3047133 A	28-02-1991
		PT 93166 A,B	31-08-1990
WO 9707198 A	27-02-1997	US 5707829 A	13-01-1998
		AU 6712396 A	18-02-1997
		AU 6768596 A	12-03-1997
		EP 0839196 A	06-05-1998
		EP 0851875 A	08-07-1998
		WO 9704097 A	06-02-1997
US 5536637 A	16-07-1996	US 5712116 A	27-01-1998